example in 31,570 human CoV-2 genomes of a substitution that enhances ACE2 binding, the CoV-2 interaction with ACE-2 was maximized from the get-go.

Therefore, the hypothesis, "If the SARS-CoV-2 (CoV-2) Spike Protein interaction with the ACE2 receptor is not maximized, then it is evidence that the interaction is the product of natural selection and not purposeful (laboratory) manipulation," is **rejected**.

The alternative hypothesis, "If the CoV-2 Spike Protein interaction with the ACE2 receptor is maximized, then it is evidence that the interaction was the product of purposeful (laboratory) manipulation," is thus **accepted**.

At the time of this writing, a new RBD mutant N501Y has been observed. It is one of the five potential mutations that could be expected to increase RBD-ACE2 affinity.

This is the first example of evidence that will not be statistically quantified but treated as a 51%.49% preponderance of the evidence adjustment. The evidence is more consistent with having been optimized by various methods used in the laboratory than with the slow natural process as seen with SARS-CoV-1, and so the conservative rule that this is consistent with a laboratory origin (51%) versus zoonotic origin (49%) will be used. There will be no confidence adjustment.

The adjusted likelihoods are shown in the following table.

Evidence or process	Zoonotic Origin (ZO)	Laboratory Origin (LO)
Starting likelihood	0.002	0.998
This is the outcome favors LO over ZO at 51% versus 49%		0.51
Impact of this evidence		Increases the likelihood of LO by 51/49 = 1.041
Impact of evidence calculation		1.041 x 0.998 = 1.039
Normalize this step of analysis	0.002/(0.002 + 1.039) = 0.002	1.039/(0.002 + 1.039) = 0.998

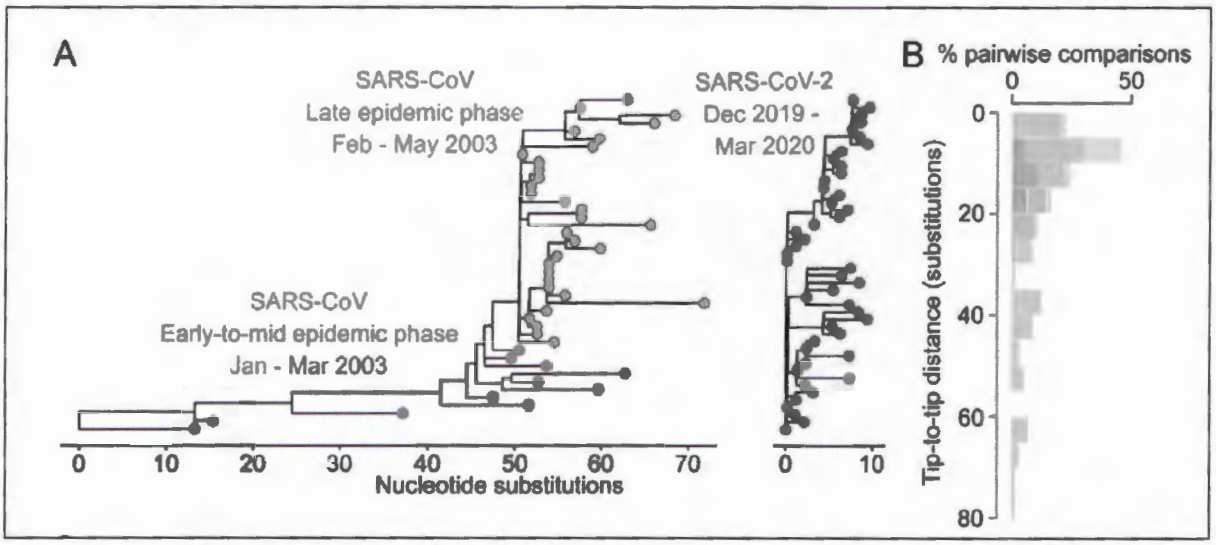
Adjusted likelihood: Zoonotic origin (0.2%), laboratory origin (99.8%).

Evidence. Whole genome comparison of human adaption of CoV-2 compared to SARS-CoV-1 is consistent with a "pre-adaption" of CoV-2 to the human host

A paper[99] entitled, "SARS-CoV-2 is well adapted for humans. What does this mean for re-emergence?" by Shing Hei Zhan, Benjamin E. Deverman, and Yujia Alina Chan states in the abstract:

"In a side-by-side comparison of evolutionary dynamics between the 2019/2020 SARS-CoV-2 and the 2003 SARS-CoV, we were surprised to find that SARS-CoV-2 resembles SARS-CoV in the late phase of the 2003 epidemic, after SARS-CoV had developed several advantageous adaptations for human transmission. Our observations suggest that **by the time SARS-CoV-2 was first detected in late 2019, it was already pre-adapted to human transmission to an extent similar to late epidemic SARS-CoV. However, no precursors or branches of evolution stemming from a less human-adapted SARS-CoV-2-like virus have been detected.** The sudden appearance of a highly infectious SARS-CoV-2 presents a major cause for concern that should motivate stronger international efforts to identify the source and prevent re-emergence in the near future. [Emphasis added.]

The following Figure from the paper best illustrates the relative SNV adaption for SARS-CoV-1 versus CoV-2.



The paper also makes a tangential comment about posterior diversity: "It would be curious if no precursors or branches of SARS-CoV-2 evolution are discovered in humans or animals."

This is another example of evidence that will not be statistically quantified. The evidence is more consistent with having been adapted by various known methods used in a laboratory than with the slow natural process as seen with SARS-CoV-1, and so the conservative rule that this is consistent with a laboratory origin (51%) versus zoonotic origin (49%) will be used. There will be no confidence adjustment.

[99] https://www.biorxiv.org/content/10.1101/2020.05.01.073262v1

The adjusted likelihoods are shown in the following table.

Evidence or process	Zoonotic Origin (ZO)	Laboratory Origin (LO)
Starting likelihood	0.002	0.998
This is the outcome favors LO over ZO at 51% versus 49%		0.51
Impact of this evidence		Increases the likelihood of LO by 51/49 = 1.041
Impact of evidence calculation		1.041 x 0.998 = 1.039
Normalize this step of analysis	0.002/(0.002 + 1.039) = 0.002	1.039/(0.002 + 1.039) = 0.998

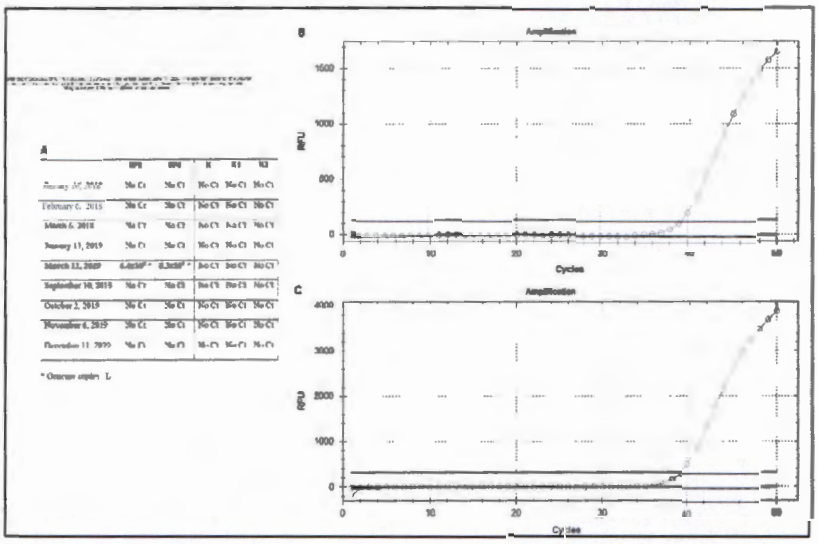
Adjusted likelihood: Zoonotic origin (0.2%), laboratory origin (99.8%).

Evidence: Evidence of CoV-2 during early 2019 in wastewater from Barcelona, Spain is a false positive artifact

A paper entitled "Sentinel surveillance of SARS-CoV-2 in wastewater anticipates the occurrence of COVID-19 cases"[100] claims CoV-2 was present in Barcelona, Spain in March 2019. Specifically, they state:

"This possibility prompted us to analyze some archival WWTP samples from January 2018 to December 2019 (Figure 2). All samples came out to be negative for the presence of SARS-CoV-2 genomes with the exception of March 12, 2019, in which both IP2 and IP4 target assays were positive. This striking finding indicates circulation of the virus in Barcelona long before the report of any COVID-19 case worldwide."

This is a false positive



As shown above from the paper, they found 43/45 runs with zero and two runs had only 600-800 CoV-2 copies/L

But the limit of detection (LoD) of their assay is 1,000,000 CoV-2/L.

According to the Promega PCR assay FDA clearance package, the Ct at the LoD is 33-34 for the N1 and N2, respectively (Table 17, page 51).[101] Here the LoD is listed as 1 RNA/μL.

In the paper the Ct is 40 or 6-7 above the LoD.

This evidence is neutral as to origin and will not be used to adjust the likelihoods. It does reduce the credibility of some of the new origin theories coming out of China.

[100] https://www.medrxiv.org/content/10.1101/2020.06.13.20129627v1.full.pdf
[101] https://twitter.com/quay_dr/status/1340572543548227585/photo/1

Evidence: WHO and Dr. Shi have spoken of the singular nature of the beginning of COVID-19

On January 23, 2020 Dr. Shi wrote in the draft of her paper: "The almost identical sequences of this virus in different patients imply a probably recent introduction in humans…"[102] By February 3, 2020, when the final version of this paper was published, this sentence had been **deleted**.[103]

On April 23, 2020 the WHO stated: "All the published genetic sequences of SARS-CoV-2 isolated from human cases are very similar. This suggests that the start of the outbreak resulted from a single point introduction in the human population around the time that the virus was first reported in humans in Wuhan, China in December 2019."[104]

The evidence, like the lack of posterior diversity and seroconversion reported earlier, is more consistent with a single introduction in a laboratory accident. This evidence will not be used to adjust probabilities but is included because it could be a form of party admissions of unfavorable facts.

[102] RaTG13 paper as a preprint
[103] RaTG13 'final' Nature paper
[104] WHO document page 2 of 12

Evidence. As documented by Drs. Daszak, Humes, and Shi, mammalian biodiversity and bat species differences between Yunnan and Hubei Provence are significant and do not support a zoonotic origin

Summary. SARS-CoV-2 is most closely related to bat coronaviruses from Yunnan, a rural province in South West China. Wuhan, where the pandemic began, is a large urban city of 11 million inhabitants in north central China. These two areas are approximately 1900 km apart.

This is the US equivalent of the difference between New York City (population 8.4 million) and the Everglades in Florida, 2000 km away. The incongruent image of a bat or intermediate host in the Everglades somehow finding its way to New York City is a clear demonstration of the difficulty in this hypothetical transmission process. Nonetheless, a strict literature-based analysis will be conducted.

If COVID-19 is a zoonotic disease it must have travelled from bats to humans or from bats to an intermediate species to humans. Therefore, an examination of mammalian biodiversity differences and commonalities between Yunnan and Wuhan might provide useful information about the intermediate host or the particular bat species.

Peter Daszak, Zhengli-li Shi and colleagues published an August 2020 paper entitled, "Origin and cross-species transmission of bat coronaviruses in China,"[105] in which they make a number of observations that are relevant to this analysis. It should be remembered that both lead authors have made multiple, strong, public statements over many months where they assert that SARS-CoV-2 is a natural virus of zoonotic origin.

Yunnan and Hubei Provinces have very dissimilar mammalian diversity

Quoting from the Methods section of the Daszak, Shi paper:

"Defining zoogeographic regions in China:

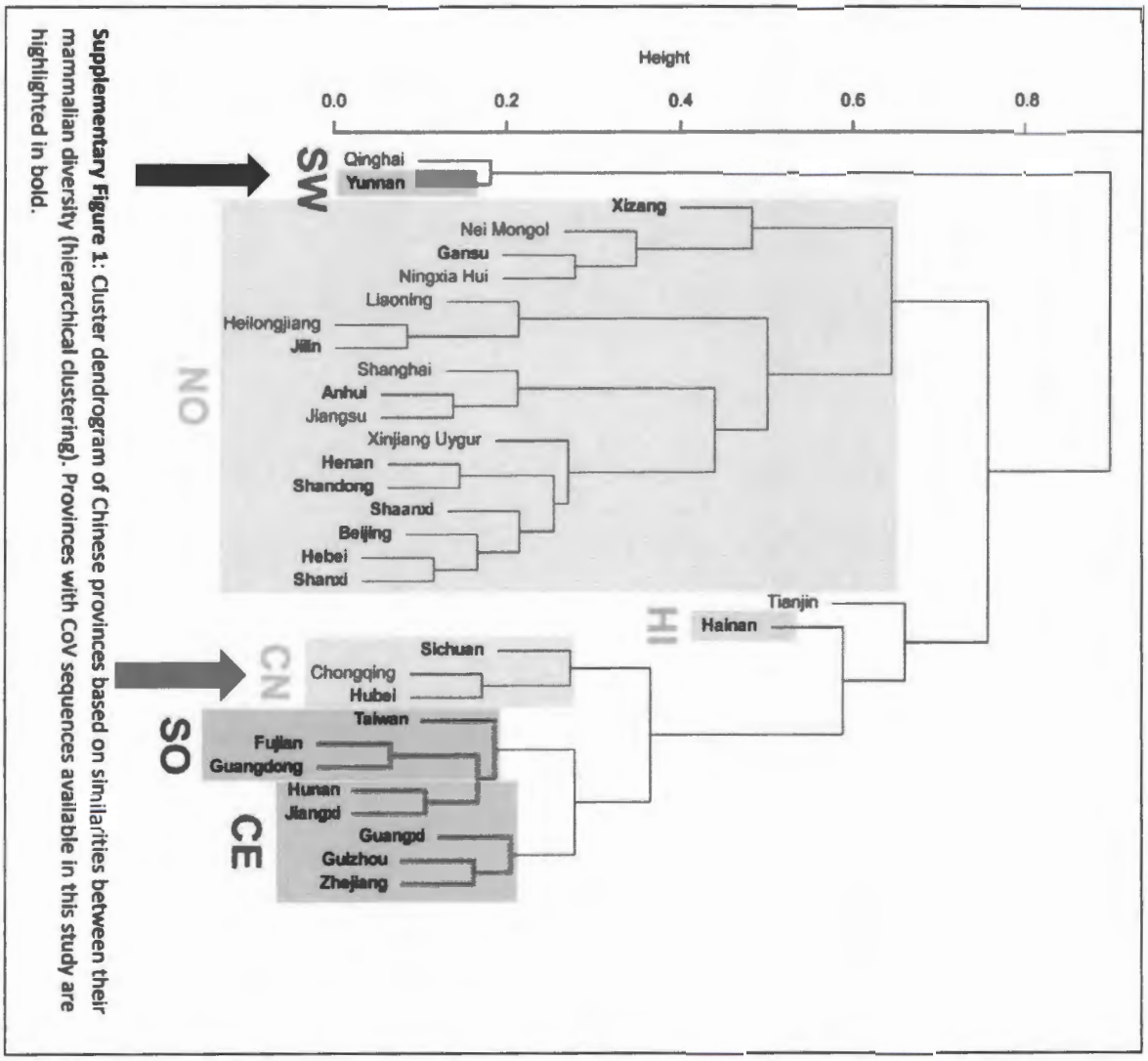
Hierarchical clustering was used to define zoogeographic regions within China by clustering provinces with similar mammalian diversity. Hierarchical cluster analysis classifies several objects into small groups based on similarities between them. To do this, we created a presence/absence matrix of all extant terrestrial mammals present in China using data from the IUCN spatial database and generated a cluster dendrogram using the function *hclust* with average method of the R package stats. Hong Kong and Macau were included within the neighboring Guangdong province. We then visually identified geographically contiguous clusters of provinces for which CoV sequences are available (Fig. 1 and Supplementary Fig. 1).

We identified six zoogeographic regions within China based on the similarity of the mammal community in these provinces: **SW (Yunnan province)**, NO (Xizang, Gansu, Jilin, Anhui, Henan, Shandong, Shaanxi, Hebei, and Shanxi provinces and Beijing municipality), **CN (Sichuan and Hubei provinces)**, CE (Guangxi, Guizhou, Hunan, Jiangxi, and Zhejiang provinces), SO (Guangdong and Fujian provinces, Hong Kong, Macau, and Taiwan), and HI.

[105] https://www.nature.com/articles/s41467-020-17687-3#Sec19

Hunan and Jiangxi, clustering with the SO provinces in our dendrogram, were included within the central region to create a geographically contiguous Central cluster (Supplementary Fig. 1). These six zoogeographic regions are very similar to the biogeographic regions traditionally recognized in China. The three β-CoV sequences from HI were included in the SO region to avoid creating a cluster with a very small number of sequences."

Below is a cluster dendrogram of Chinese provinces based on similarities between their mammalian diversity (hierarchical clustering). Provinces with CoV sequences available in this study are highlighted in bold.



Supplementary Figure 1: Cluster dendrogram of Chinese provinces based on similarities between their mammalian diversity (hierarchical clustering). Provinces with CoV sequences available in this study are highlighted in bold.

The y-axis height is a measure of the biodiversity with 1.0 being complete similarity and 0.0 being no similarity. As expected for the geography and location of the two provinces, Yunnan (red arrow above) and Hubei (green arrow above) have a height score of about 0.1, with seven branches and six nodes separating them. This is close to the biggest different in mammalian biodiversity of any two locations in all of China.

In conclusion, Daszak and Shi et al. demonstrate that the mammalian biodiversity between Yunnan and Hubei is very significant, reducing the options for a common intermediate host to be the natural conduit between bats and humans.

Shi, Humes, and Daszak statement: "SARS-CoV-2 is likely derived from a clade of viruses originating in horseshoe bats (*Rhinolophus* spp.). The geographic location of this origin appears to be Yunnan province."

This evidence will not be statistically quantified. The evidence reduces the biodiversity overlap needed to create a common intermediate species between the two provinces, and so the conservative rule that this is consistent with a laboratory origin (51%) versus zoonotic origin (49%) will be used. There will be no subjective discount factor adjustment.

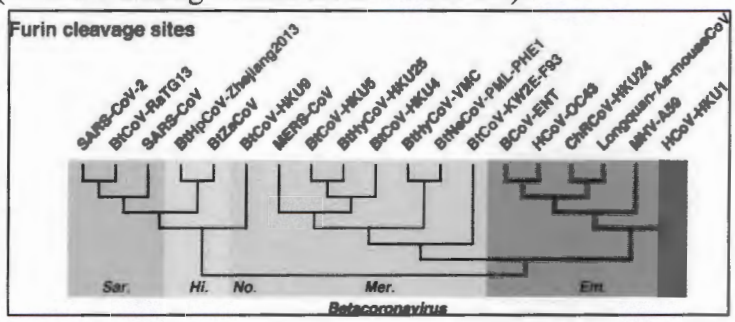
Evidence or process	Zoonotic Origin (ZO)	Laboratory Origin (LO)
Starting likelihood	0.002	0.998
This is the outcome favors LO over ZO at 51% versus 49%		0.51
Impact of this evidence		Increases the likelihood of LO by 51/49 = 1.041
Impact of evidence calculation		1.041 x 0.998 = 1.039
Normalize this step of analysis	0.002/(0.002 + 1.039) = 0.002	1.039/(0.002 + 1.039) = 0.998

Because of the rule on the use of significant figures, the likelihood does not change.

Adjusted likelihood: Zoonotic origin (0.2%), laboratory origin (99.8%).

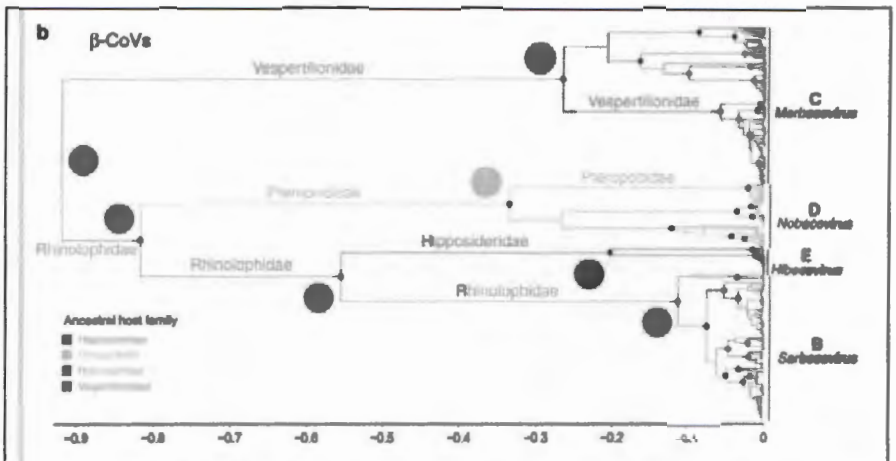
Evidence: The ancestor of SARS-CoV-2 can hypothetically only obtain a furin site by recombination outside of the sarbecovirus subgenera but there is strong evidence that coronavirus recombination is largely limited to the clade level, with limited evidence of sub-genera or genera recombination

- SARS-CoV-2 is a beta coronavirus, subgenera sarbecovirus and is the only sarbecovirus with a furin site.[106]
- Furin sites can be found in either alpha or gamma coronaviruses or the other beta coronavirus subgenera. The following Figure from reference 66 shows examples of such coronaviruses (furin containing viruses are shown in red):



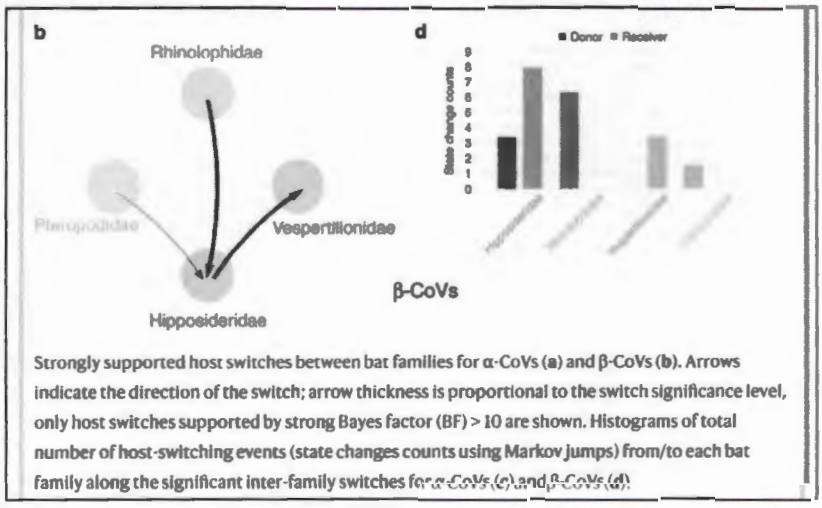
- To acquire a furin site in nature would require a co-infection between the CoV-2 sarbecovirus ancestor and a furin-containing non-sarbecovirus as shown above.
- However, there is no evidence of recombination in coronaviruses at either the genus level or the subgenus level; only at the clade level.[107][108]
- There is also evidence from Daszak and Shi that within the subgenera of the beta coronaviruses, there is bat host specificity. So, each subgenera of coronaviruses has a preferred bat host species. This reduces the opportunities for a co-host event to permit recombination.[109] The phylogeny below shows the problem of host incompatibility for beta coronaviruses:

[106] https://www.sciencedirect.com/science/article/pii/S1873506120304165#f0015
[107] file:///C:/Users/Steven%20Quay/Desktop/journal.pgen.1009272.pdf
[108] https://academic.oup.com/mbe/advance-article/doi/10.1093/molbev/msaa281/5955840
[109] https://www.nature.com/articles/s41467-020-17687-3#Sec2



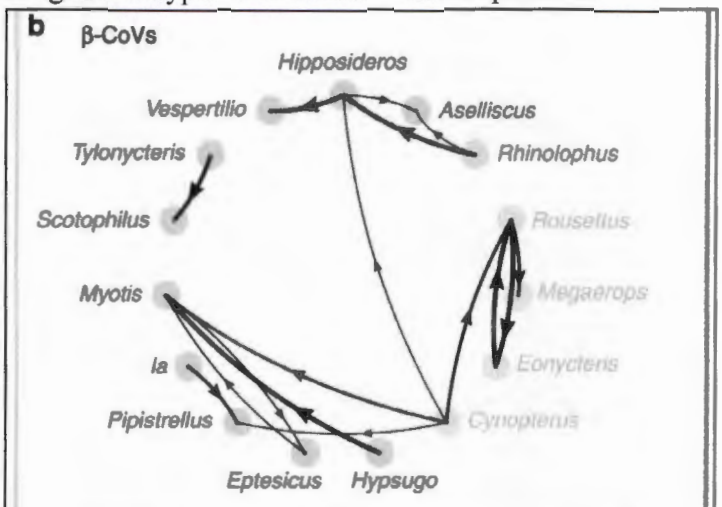
α-CoV (a) and β-CoV (b) maximum clade credibility annotated trees using complete datasets of *RdRp* sequences and bat host family as discrete character state. Pie charts located at the root and close to the deepest nodes show the state posterior probabilities for each bat family. Branch color correspond to the inferred ancestral family with the highest probability. Branch lengths are scaled according to relative time units (clock rate = 1.0). Well-supported nodes (posterior probability > 0.95) are indicated with a black dot. The ICTV approved CoV subgenera were highlighted: *Rhinacovirus* (L1), *Decacovirus* (L2), *Myotacovirus* (L3), *Pedacovirus* (L5), *Nyctacovirus* (L6), *Minunacovirus* (L7), and an unidentified lineage (L4) for α-CoVs; and *Merbecovirus* (Lineage C), *Nobecovirus* (lineage D), *Hibecovirus* (lineage E), and *Sarbecovirus* (Lineage B) for β-CoVs.

- Daszak and Shi also identified preferred directions of host switching. Since RaTG13, the closest coronavirus to SARS-CoV-2, is most closely related to viruses with bat hosts from the family, Rhinolophidae, it would be reasonable to expect furin-containing viruses from other bat hosts to migrate into Rhinolophidae, recombine by methods which have not been identified, and then the furin-containing sarbecovirus could evolve into the ancestor of SARS-CoV-2. Unexpectedly, Daszak et al. found host migration for the Rhinolophidae bats only outward and not inward, as required by the above, admittedly, convoluted process. The data Figure is shown here:



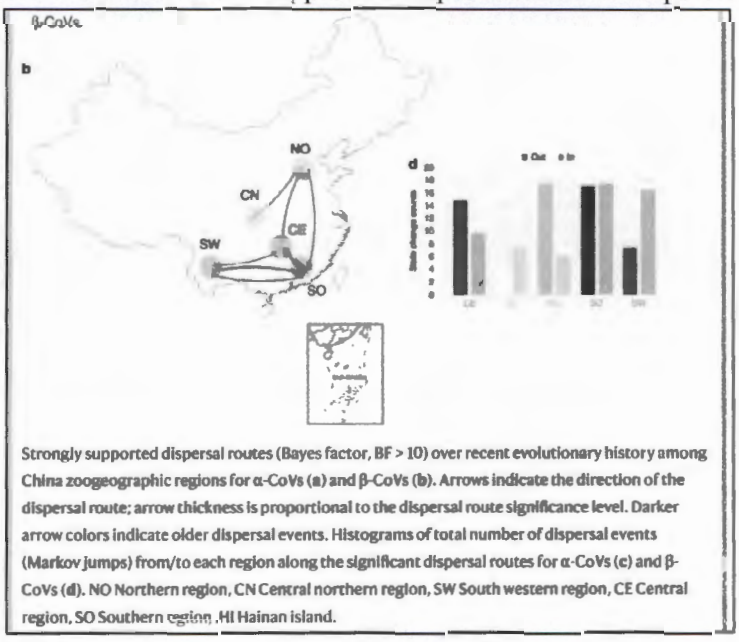
Strongly supported host switches between bat families for α-CoVs (a) and β-CoVs (b). Arrows indicate the direction of the switch; arrow thickness is proportional to the switch significance level, only host switches supported by strong Bayes factor (BF) > 10 are shown. Histograms of total number of host-switching events (state changes counts using Markov jumps) from/to each bat family along the significant inter-family switches for α-CoVs (c) and β-CoVs (d).

- Daszak and Shi also observed outward host switches from *Rhinolophus* at the genera level as well, also against a hypothesis for furin-site acquisition:



Strongly supported host switches between bat genera for α-CoVs (a) and β-CoVs (b) and their significance level (Bayes factor, BF). Only host switches supported by strong BF values >10 are shown. Line thickness is proportional to the switch significance level. Red lines correspond to host switches among bat genera belonging to different families, and black lines correspond to host switches among bat genera from the same family. Arrows indicate the direction of the switch. Genus names are colored according to the family they belong to using the same colors as in Figs. 2 and 3.

- Finally, this paper by Daszak and Shi states: "We used our Bayesian discrete phylogeographic model with zoogeographic regions as character states to reconstruct the spatiotemporal dynamics of CoV dispersal in China." If SARS-CoV-2 began in Yunnan and first crossed over into humans in Wuhan, this analysis should support a northernly spatiotemporal dispersal of beta coronaviruses. Unfortunately, Daszak and Shi cannot catch a break; their own data do not support the expected route of dispersion:



Strongly supported dispersal routes (Bayes factor, BF > 10) over recent evolutionary history among China zoogeographic regions for α-CoVs (a) and β-CoVs (b). Arrows indicate the direction of the dispersal route; arrow thickness is proportional to the dispersal route significance level. Darker arrow colors indicate older dispersal events. Histograms of total number of dispersal events (Markov jumps) from/to each region along the significant dispersal routes for α-CoVs (c) and β-CoVs (d). NO Northern region, CN Central northern region, SW South western region, CE Central region, SO Southern region ,HI Hainan island.

As shown in the above Figure the only dispersal routes into Wuhan, which is in the CN region, are from the northern region. And the northern region has no inward dispersals from the SW, southwest region, where Yunnan and the origin of the ancestor of SARS-CoV-2, is located.

- Independent evidence documents that Hubei province does not have the bat species needed for SARS-CoV-2 reservoir host[110]

While statistical models of this data could be interesting and informative for general research about future spillovers, this is evidence will not be statistically quantified for this analysis. The evidence reduces the opportunities for subgenera co-infection and furin-site recombination into the CoV-2 ancestor and so the conservative rule that this is less consistent with a zoonotic origin (49%) versus laboratory origin (51%) will be used. There will be no subjective discount factor adjustment.

The results from the calculations are shown below.

Evidence or process	Zoonotic Origin (ZO)	Laboratory Origin (LO)
Starting likelihood	0.002	0.998
This is the outcome favors LO over ZO at 51% versus 49%		0.51
Impact of this evidence		Increases the likelihood of LO by 51/49 = 1.041
Impact of evidence calculation		1.041 x 0.998 = 1.039
Normalize this step of analysis	0.002/(0.002 + 1.039) = 0.002	1.039/(0.002 + 1.039) = 0.998

Adjusted likelihood: Zoonotic origin (0.2%), laboratory origin (99.8%).

[110] file:///C:/Users/Steven%20Quay/Desktop/Zhangetal2009.pdf

Evidence: Of 410 vertebrate species tested for affinity to CoV-2 Spike Protein binding domain, primate ACE2 receptor, including human and VERO monkey cells, are the best at binding and bat species ACE2 are the worse, making direct bat-to-human host jumping extremely unlikely

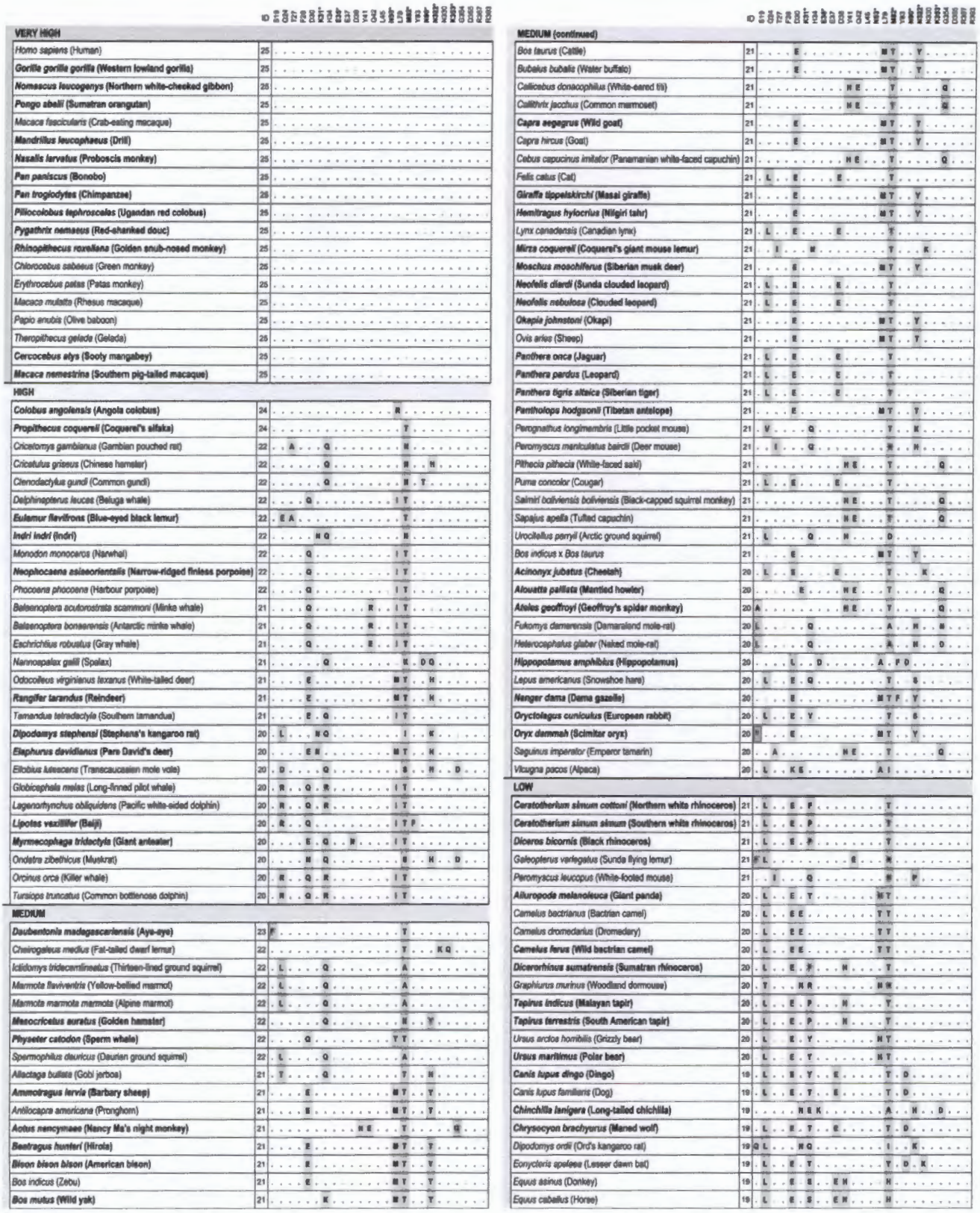
- An examination of the ACE2 receptor binding domain amino acid sequences and their suitability for interacting with SARS-CoV-2 was performed in 410 vertebrates, including 252 mammals.[111]
- A five-category binding score was developed based on the conservation properties of 25 amino acids important for the binding between ACE2 and the SARS-CoV-2 spike protein.
- Only mammals fell into the medium to very high categories and only primates scored 25/25 for binding.
- This implies that SARS-CoV-2 is optimized for human ACE2-bearing cells from the first introduction into the human population, an observation that contradicts a zoonotic origin.
- It also suggests that other primates may be the proximate species from which SARS-CoV-2 entered the human population.
- Both VERO monkey kidney cells and ACE2 humanized mice would quality as an intermediate species by this criterion.
- Surprisingly, "all chiropterans (bats) scored low (n = 8) or very low (n = 29), including the Chinese rufous horseshoe bat, from which a coronavirus (SARSr-CoV ZC45) related to SARS-CoV-2 was identified."
- This is evidence that bats are probably not a reservoir host for SARS-CoV-2.
- A separate study observed: "Severe acute respiratory syndrome coronavirus 2 did not replicate efficiently in 13 bat cell lines."[112]
- The following two Tables are taken from the paper and are organized according to ACE2 SARS-CoV-2 affinity, from highest to lowest:

[111] https://www.pnas.org/content/117/36/22311
[112] https://wwwnc.cdc.gov/eid/article/26/12/20-2308_article

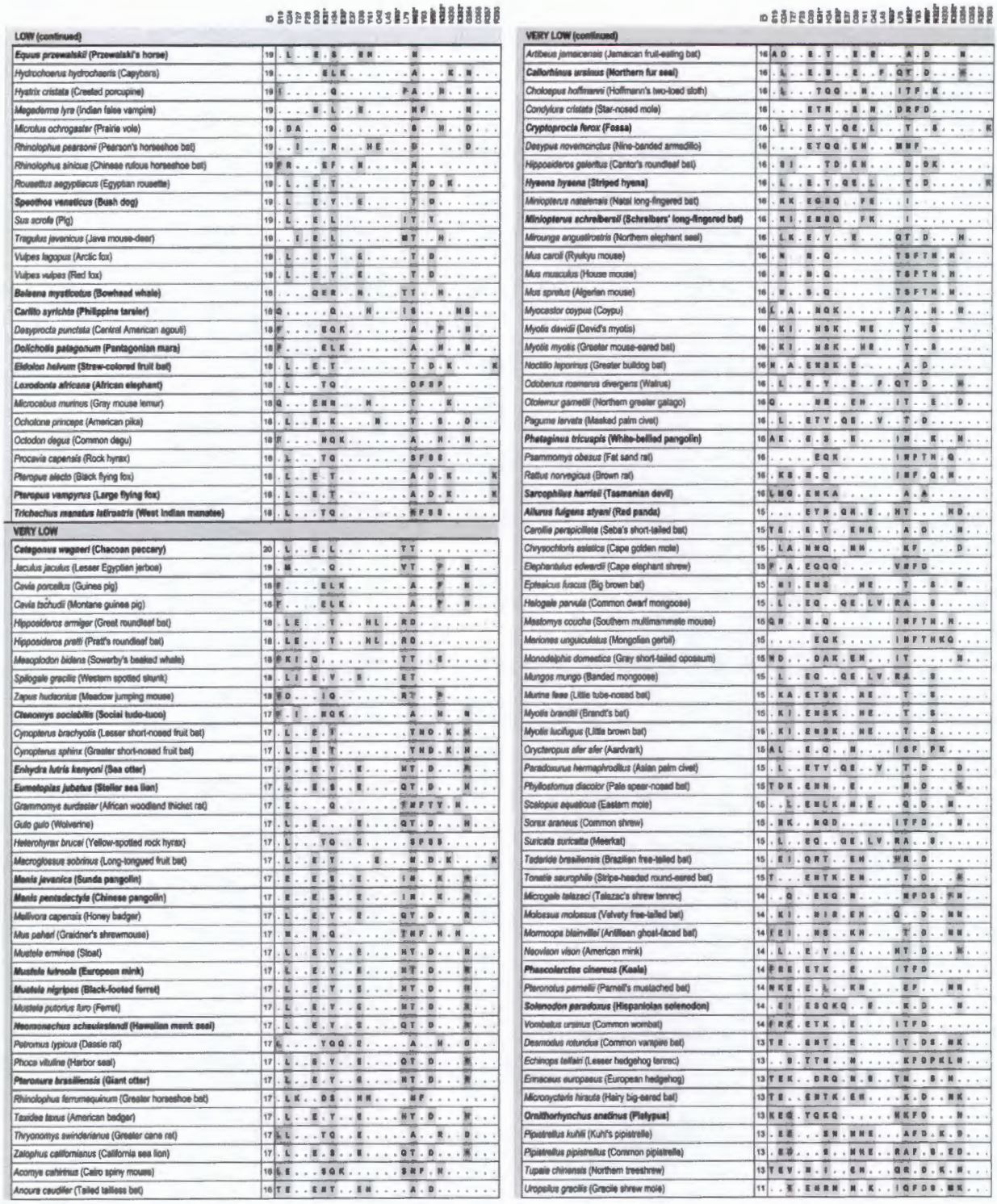
Bayesian Analysis of SARS-CoV-2 Origin

Steven C. Quay, MD, PhD

29 January 2021



Bayesian Analysis of SARS-CoV-2 Origin
Steven C. Quay, MD, PhD **29 January 2021**



LOW (continued)	ID
***Equus przewalskii* (Przewalski's horse)**	19
Hydrochoerus hydrochaeris (Capybara)	19
Hystrix cristata (Crested porcupine)	19
Megaderma lyra (Indian false vampire)	19
Microtus ochrogaster (Prairie vole)	19
Rhinolophus pearsoni (Pearson's horseshoe bat)	19
Rhinolophus sinicus (Chinese rufous horseshoe bat)	19
Rousettus aegyptiacus (Egyptian rousette)	19
***Speothos venaticus* (Bush dog)**	19
Sus scrofa (Pig)	19
Tragulus javanicus (Java mouse-deer)	19
Vulpes lagopus (Arctic fox)	19
Vulpes vulpes (Red fox)	19
***Balaena mysticetus* (Bowhead whale)**	18
***Carlito syrichta* (Philippine tarsier)**	18
Dasyprocta punctata (Central American agouti)	18
***Dolichotis patagonum* (Pantagonian mara)**	18
***Eidolon helvum* (Straw-colored fruit bat)**	19
***Loxodonta africana* (African elephant)**	18
Microcebus murinus (Gray mouse lemur)	18
Ochotona princeps (American pika)	18
Octodon degus (Common degu)	18
Procavia capensis (Rock hyrax)	18
Pteropus alecto (Black flying fox)	18
***Pteropus vampyrus* (Large flying fox)**	18
***Trichechus manatus latirostris* (West Indian manatee)**	18
VERY LOW	
***Catagonus wagneri* (Chacoan peccary)**	20
Jaculus jaculus (Lesser Egyptian jerboa)	19
Cavia porcellus (Guinea pig)	18
Cavia tschudii (Montane guinea pig)	18
Hipposideros armiger (Great roundleaf bat)	18
Hipposideros pratti (Pratt's roundleaf bat)	18
Mesoplodon bidens (Sowerby's beaked whale)	18
Spilogale gracilis (Western spotted skunk)	18
Zapus hudsonius (Meadow jumping mouse)	18
***Ctenomys sociabilis* (Social tuco-tuco)**	17
Cynopterus brachyotis (Lesser short-nosed fruit bat)	17
Cynopterus sphinx (Greater short-nosed fruit bat)	17
***Enhydra lutris kenyoni* (Sea otter)**	17
***Eumetopias jubatus* (Steller sea lion)**	17
Grammomys surdaster (African woodland thicket rat)	17
Gulo gulo (Wolverine)	17
Heterohyrax brucei (Yellow-spotted rock hyrax)	17
Macroglossus sobrinus (Long-tongued fruit bat)	17
***Manis javanica* (Sunda pangolin)**	17
***Manis pentadactyla* (Chinese pangolin)**	17
Mellivora capensis (Honey badger)	17
Mus pahari (Gnaidner's shrewmouse)	17
Mustela erminea (Stoat)	17
***Mustela lutreola* (European mink)**	17
***Mustela nigripes* (Black-footed ferret)**	17
Mustela putorius furo (Ferret)	17
***Neomonachus schauinslandi* (Hawaiian monk seal)**	17
Petromus typicus (Dassie rat)	17
Phoca vitulina (Harbor seal)	17
***Pteronura brasiliensis* (Giant otter)**	17
Rhinolophus ferrumequinum (Greater horseshoe bat)	17
Taxidea taxus (American badger)	17
Thryonomys swinderianus (Greater cane rat)	17
Zalophus californianus (California sea lion)	17
Acomys cahirinus (Cairo spiny mouse)	16
Anoura caudifer (Tailed tailless bat)	16

VERY LOW (continued)	ID
Artibeus jamaicensis (Jamaican fruit-eating bat)	16
***Callorhinus ursinus* (Northern fur seal)**	16
Choloepus hoffmanni (Hoffmann's two-toed sloth)	16
Condylura cristata (Star-nosed mole)	16
***Cryptoprocta ferox* (Fossa)**	16
Dasypus novemcinctus (Nine-banded armadillo)	16
Hipposideros galeritus (Cantor's roundleaf bat)	16
***Hyaena hyaena* (Striped hyena)**	16
Miniopterus natalensis (Natal long-fingered bat)	16
***Miniopterus schreibersii* (Schreibers' long-fingered bat)**	16
Mirounga angustirostris (Northern elephant seal)	16
Mus caroli (Ryukyu mouse)	16
Mus musculus (House mouse)	16
Mus spretus (Algerian mouse)	16
Myocastor coypus (Coypu)	16
Myotis davidii (David's myotis)	16
Myotis myotis (Greater mouse-eared bat)	16
Noctilio leporinus (Greater bulldog bat)	16
Odobenus rosmarus divergens (Walrus)	16
Otolemur garnettii (Northern greater galago)	16
Paguma larvata (Masked palm civet)	16
***Phataginus tricuspis* (White-bellied pangolin)**	16
Psammomys obesus (Fat sand rat)	16
Rattus norvegicus (Brown rat)	16
***Sarcophilus harrisii* (Tasmanian devil)**	16
***Ailurus fulgens styani* (Red panda)**	15
Carollia perspicillata (Seba's short-tailed bat)	15
Chrysochloris asiatica (Cape golden mole)	15
Elephantulus edwardii (Cape elephant shrew)	15
Eptesicus fuscus (Big brown bat)	15
Helogale parvula (Common dwarf mongoose)	15
Mastomys coucha (Southern multimammate mouse)	15
Meriones unguiculatus (Mongolian gerbil)	15
Monodelphis domestica (Gray short-tailed opossum)	15
Mungos mungo (Banded mongoose)	15
Murina feae (Little tube-nosed bat)	15
Myotis brandtii (Brandt's bat)	15
Myotis lucifugus (Little brown bat)	15
Orycteropus afer afer (Aardvark)	15
Paradoxurus hermaphroditus (Asian palm civet)	15
Phyllostomus discolor (Pale spear-nosed bat)	15
Scalopus aquaticus (Eastern mole)	15
Sorex araneus (Common shrew)	15
Suricata suricatta (Meerkat)	15
Tadarida brasiliensis (Brazilian free-tailed bat)	15
Tonatia saurophila (Stripe-headed round-eared bat)	15
Microgale talazaci (Talazac's shrew tenrec)	14
Molossus molossus (Velvety free-tailed bat)	14
Mormoops blainvillei (Antillean ghost-faced bat)	14
Neovison vison (American mink)	14
***Phascolarctos cinereus* (Koala)**	14
Pteronotus parnellii (Parnell's mustached bat)	14
***Solenodon paradoxus* (Hispaniolan solenodon)**	14
Vombatus ursinus (Common wombat)	14
Desmodus rotundus (Common vampire bat)	13
Echinops telfairi (Lesser hedgehog tenrec)	13
Erinaceus europaeus (European hedgehog)	13
Micronycteris hirsuta (Hairy big-eared bat)	13
***Ornithorhynchus anatinus* (Platypus)**	13
Pipistrellus kuhlii (Kuhl's pipistrelle)	13
Pipistrellus pipistrellus (Common pipistrelle)	13
Tupaia chinensis (Northern treeshrew)	13
Uropsilus gracilis (Gracile shrew mole)	11

While statistical models of this data could be interesting and informative, this is evidence will not be statistically quantified for this analysis. The evidence is another way of looking at the pre-adapted state of the CoV-2 for humans and suggests that primate animals, monkey cell cultures like the VERO cell, and humanized mice could be likely laboratory models that were used by the

WIV in GoF research. This will contribute a 51%/49% contribution in favor of laboratory compared to zoonotic origin. There will be no subjective discount factor adjustment.

The results from the calculations are shown below.

Evidence or process	Zoonotic Origin (ZO)	Laboratory Origin (LO)
Starting likelihood	0.002	0.998
This is the outcome favors LO over ZO at 51% versus 49%		0.51
Impact of this evidence		Increases the likelihood of LO by 51/49 = 1.041
Impact of evidence calculation		1.041 x 0.998 = 1.039
Normalize this step of analysis	0.002/(0.002 + 1.039) = 0.002	1.039/(0.002 + 1.039) = 0.998

Adjusted likelihood: Zoonotic origin (0.2%), laboratory origin (99.8%).

Evidence: Did a Review of Samples Collected from a Mineshaft Cause the COVID-19 Pandemic?[113]

Abstract. The origin of the COVID-19 pandemic caused by SARS-CoV-2 has been hotly debated. Proponents of the natural spillover theory allege that the virus jumped species, possibly via an intermediary host, to cross over to humans via the wildlife trade or by other means. Proponents of a rival theory claim that the virus escaped from a laboratory in Wuhan. This research presents circumstantial evidence of a transmission route via a late 2019 review of samples collected from a mineshaft in Mojiang, Yunnan Province, China. It examines the activity at the Wuhan Institute of Virology in late 2019, when samples from a mineshaft associated with a suspected SARS outbreak were being reviewed. It proposes that spillover occurred during this review of samples including of a virus (BtCoV/4991) only 1% different to SARS-CoV-2 in its RNA-dependent RNA polymerase (RdRp).

It is a meticulous sourced analysis. It purposely avoids the question of whether SARS-CoV-2 was being grown or manipulated in the laboratory, but only addresses the evidence that events in the fall of 2019 are consistent with a laboratory accident.

This will not be used to adjust the likelihoods.

Current likelihood: Zoonotic origin (0.2%), laboratory origin (99.8%).

[113] https://zenodo.org/record/4029545#.X-x_f9gzbOg. Author anonymous. A meticulously documented analysis that concludes an accident occurred at the Wuhan Institute of Virology during the fall of 2019. Includes many primary documents from Mandarin. No direct evidence of 'what' was the nature of the accident or if it was SARS-CoV-2.

Evidence: The Hunan market was not the source of SARS-CoV-2

From the WHO Terms of Reference for the investigation of the origin of SARS-CoV-2:[114]

"The Huanan wholesale market is a large market (653 stalls and more than 1180 employees) mainly supplying seafood products but also fresh fruits and vegetables, meat, and live animals. In late December 2019, 10 stall operators were trading live wild animals including chipmunks, foxes, racoons, wild boar, giant salamanders, hedgehogs, sika deer, and many others. Farmed, wild and domestic animals were also traded at the market including snakes, frogs, quails, bamboo rats, rabbits, crocodiles, and badgers. The market was closed on 1 January 2020, and several investigations followed, including environmental sampling, as well as sampling of frozen animal carcasses at the market. **Of the 336 samples collected from animals, none were PCR positive for SARS-CoV-2**, whereas 69 out of 842 environmental samples were positive by PCR for SARS-CoV-2. Sixty- one of those (88%) were from the western wing of the market. Of these, 22 samples were from 8 different drains and sewage, and 3 viruses were isolated, sequenced and shared on GISAID. These were virtually identical to the patient samples collected at the same time (>99.9 % homology)."

For contrast, with SARS-CoV-1 91 civets & 15 raccoon dogs in wet markets were tested with 106/106, 100% positive.[115]

This will not be used to adjust the likelihoods.

Current likelihood: Zoonotic origin (0.2%), laboratory origin (99.8%).

[114] https://drive.google.com/file/d/1rx0W2efbE0R1Aq-lALWTqD22VsWbTlO-/view
[115] https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1212604/

Evidence: Analysis of the hospital of admission for COVID-19 patients during December 2019 places "ground zero" for the outbreak somewhere along Line 2 of the Wuhan Metro System.

Line 2 carries one million people per day and services the Wuhan Institute of Virology, the Hunan Seafood Market, the high-speed rail system, and the Wuhan International Airport

A preprint manuscript[116] reported that the earliest genomic cluster of SARS-CoV-2 patients is a group of four individuals associated with the General Hospital of Central Theater Command of People's Liberation Army (PLA) of China in Wuhan. This cluster contains the "Founder Patients" of both Clade A and Clade B, from which every SARS-CoV-2 coronavirus that has infected every patient with COVID-19 anywhere in the world has arisen.

The PLA Hospital is about one mile from the Wuhan Institute of Virology (WIV) and the closest hospital to WIV. Both the PLA Hospital and WIV are serviced by Line 2 of the Wuhan Metro System. The Hunan Seafood Market is also located adjacent to Line 2. All patients between December 1st, 2019 and early January 2020 were first seen at hospitals that also are serviced by Line 2 of the Metro system.

With 40 hospitals located near seven of the nine Metro Lines, the likelihood that all early patients were seen at hospitals only near Line 2 by chance is about 1 in 68,500 (p-value = 0.0000146). The inference then would be that the early spread of SARS-CoV-2 was through human-to human transmission on Line 2.

Line 2 carries one million passengers per day and assuming most are round trip business workers going to and from work in the morning and evening, represents 500,000 riders or about 5% of the Wuhan population. A very recent publication determined that, in fact, 500,000 residents of Wuhan contracted COVID-19, a ten-fold upper estimate.[117] The coincidence of my prediction that 500,000 riders on Line 2 were likely exposed to SARS-CoV-2 in late 2019 and the recent admission from Chinese CDC that Wuhan had 500,000 COVID-19 cases is duly noted!

Line 2 connects to all eight other lines of the Wuhan Metro System (1, 3, 4, 6, 7, 8, 11, and Yanglu) facilitating rapid spread in Wuhan and Hubei Province, and also services both the high-speed rail station (Hankou Railway Station), facilitating rapid spread throughout China, and the Wuhan International Airport (Tianhe International Airport), facilitating rapid spread throughout Asia, Europe, and to the United States. In fact, direct human-to-human spread from the Reference Sequence patient to patients around the world is suggested by an unexpectedly reduced genome base substitution rate seen in patient specimens in cities with direct flights from Wuhan.

---

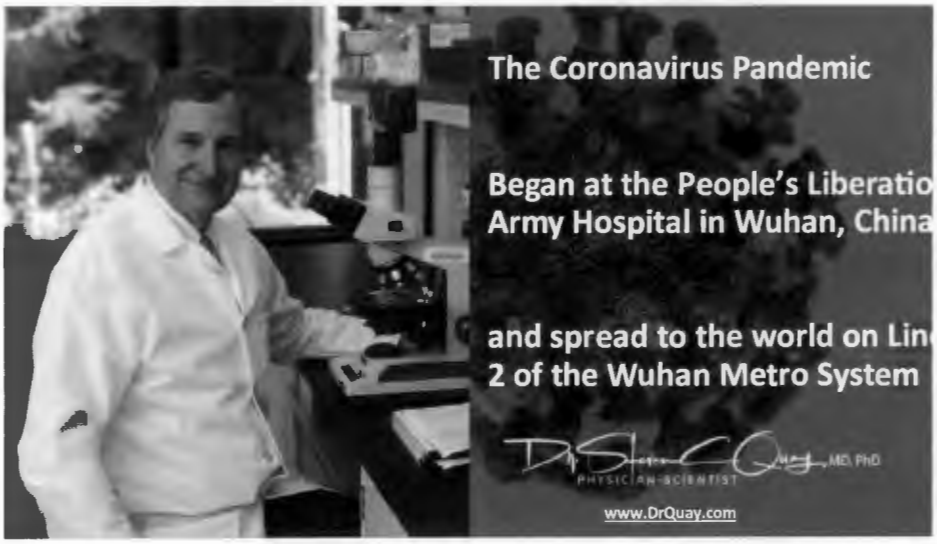
[116] https://zenodo.org/record/4119263#.X-rszNgzbOg

[117] https://mp.weixin.qq.com/s/LXTfDmsQLf3qZnu_S_MxcA ; https://thehill.com/policy/international/china/531935-study-shows-wuhan-coronavirus-cases-may-have-been-10-times-higher

In a separate paper by Quay and Dr. Martin Lee, Adjunct Professor of Statistics, UCLA, from May 2020, now accepted for publication in *Epidemics*,[118] the authors provide evidence that COVID-19 was appearing in California as early as the first week of 2020. This is likely due to direct flights connecting Line 2 to the Wuhan airport and then to San Francisco.

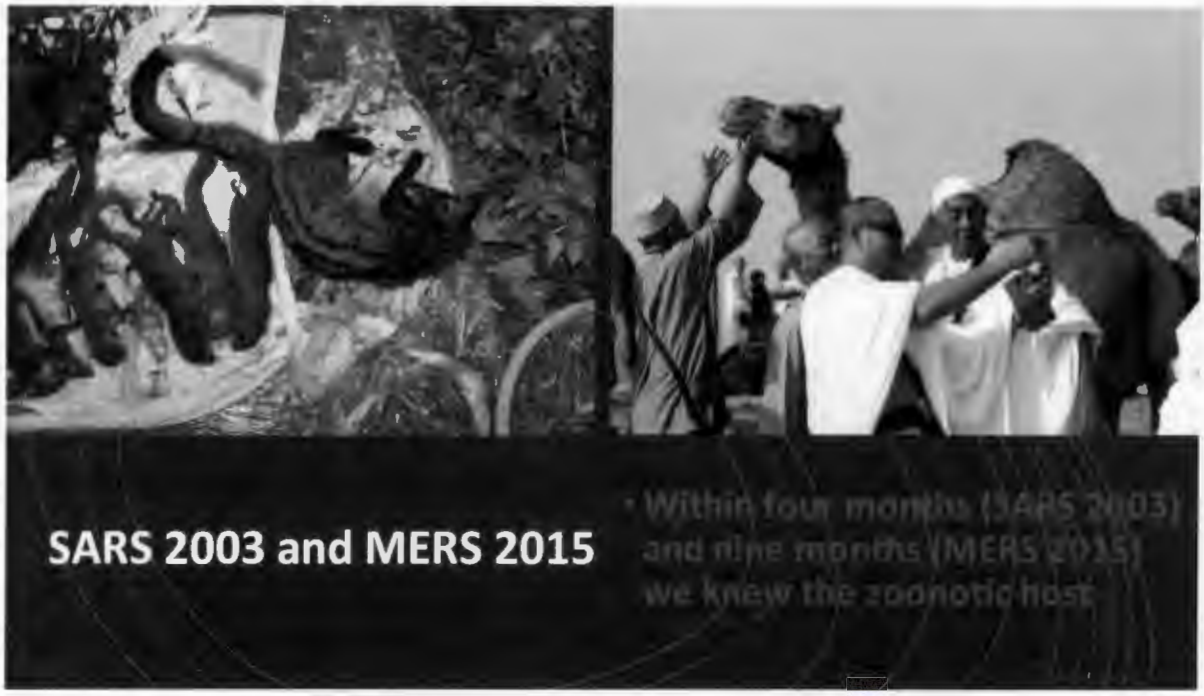
In conclusion, Line 2 of the Wuhan Metro System services the PLA Hospital with the first genomic cluster of patients with COVID-19, the hospitals where patients first went in December 2019 and early January 2020 and is the likely conduit for human-to-human spread throughout Wuhan, China, and the world.

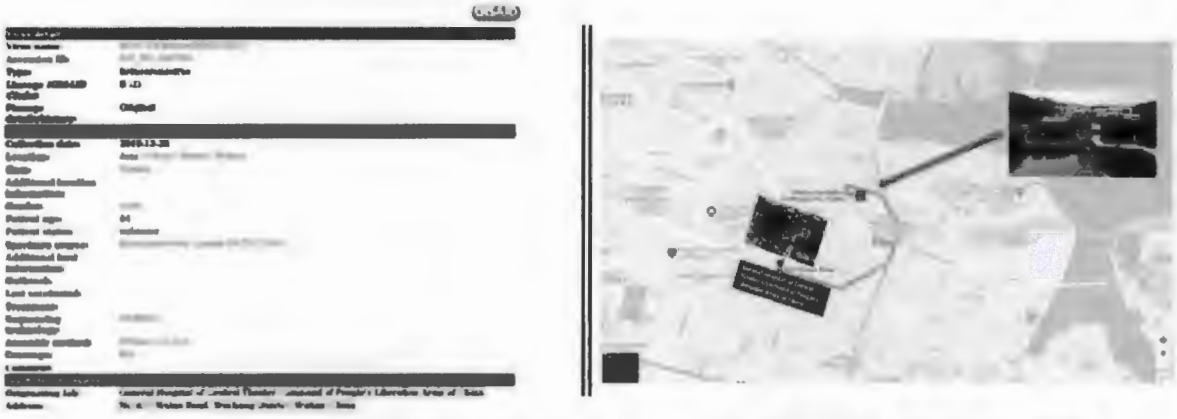
The following slide overview provides a visual analysis of this evidence:



Zoonotic Origin	Laboratory Origin
Hunan Seafood Market	Wuhan Institute of Virology (WIV); Wuhan Center for Disease Control and Prevention (CDC)

[118] https://www.researchgate.net/publication/341742303_COVID-19_May_Have_Have_Reached_United_States_in_January_2020_05272020


SARS 2003 and MERS 2015
Within four months (SARS 2003) and nine months (MERS 2015) we knew the zoonotic host


GISAID Database
Earliest cases at the PLA Hospital

Position in RS	Bat-SL-CoVZC45	Bat-SL-CoVZXC21	RaTG13	PLA-4	PLA-3	PLA-2	Hu-1 Ref seq	PLA-1	GISAID #1
5' UTR	1-5 missing	1-5 missing	1-15 missing	1-16 missing	1-20 missing	1-36 missing	Intact	1-25 missing	Intact
3778	A	A	A	A	A	A	A	A	G
6968	T	T	C	C	C	C	C	A	C
8782	T	T	T	T	C	C	C	C	C
8987	T	T	T	T	T	T	T	T	A
11764	T	T	T	T	T	NA - Note 1	T	A	T
28144	C	C	C	C	T	T	T	T	T
3' UTR	last 4 poly-A missing	last 4 poly-A missing	last 13 poly-A missing	last 15 poly-A missing	last 15 poly-A missing	NA - Note 1	Intact	last 12 poly-A missing	last 4 poly-A missing
Genome length	29802	29732	29855	29872	29868	NA - Note 1	29903	29866	29899

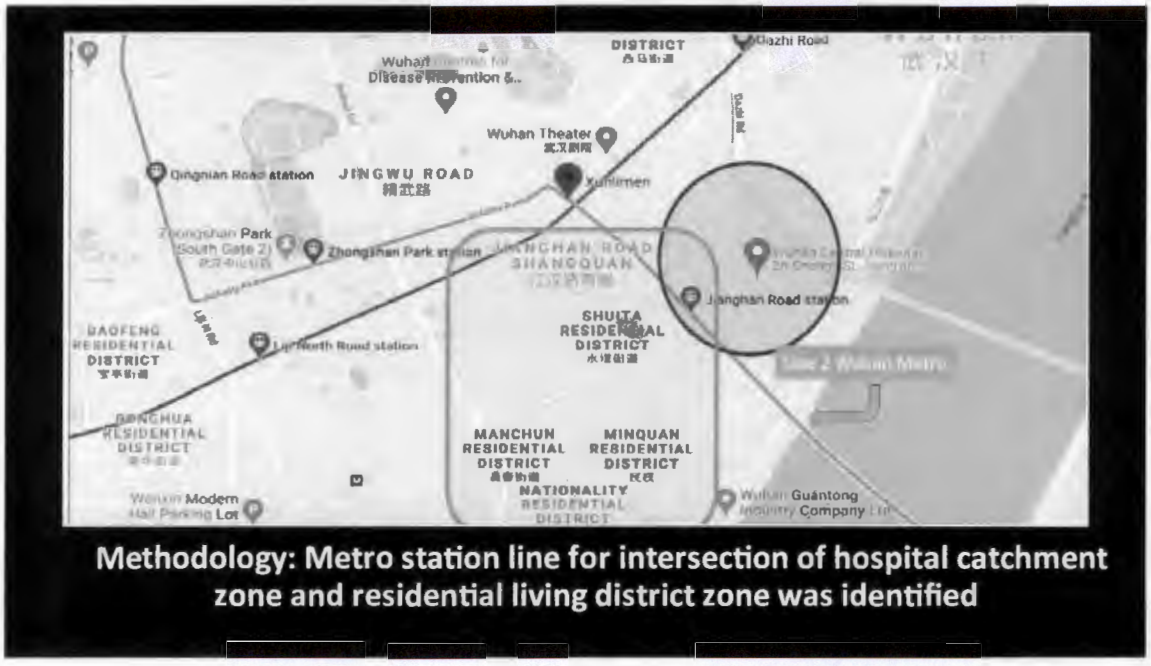
Clade A SNPs Clade B SNPs Non-RaTG13 DNPs

Note 1 - GISAID record: "Long stretches of NNNs (34.45% of overall sequence). Gap of 13 nucleotide(s) found at refpos 26171 (FRAMESHIFT). Gap of 13 nucleotides when compared to the reference sequence. 0.40% Unique Mutations."

The PLA patient cluster

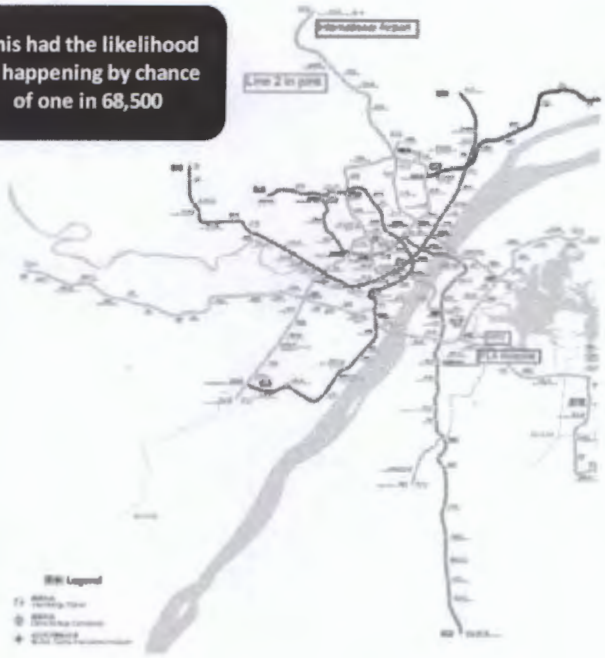
- PLA-4 is genetically the closest human infection to the three closest bat viruses
- The four PLA patients have the close sequence pattern usually seen only in family transmissions


The Wuhan Metro System


DISTRICT
Qiaozhi Road
Wuhan Theater
Qinghan Road station
JINGWU ROAD
Xuhanmen
Zhongshan Park
Zhongshan Park station
Jianghan Road station
SHULTA RESIDENTIAL DISTRICT
BAOFENG RESIDENTIAL DISTRICT
Line 2 Wuhan Metro
MANCHUN RESIDENTIAL DISTRICT
MINQUAN RESIDENTIAL DISTRICT
NATIONALITY RESIDENTIAL DISTRICT
Methodology: Metro station line for intersection of hospital catchment zone and residential living district zone was identified

All COVID patients from Dec 1 to early Jan were admitted to hospitals on Metro Line 2

This had the likelihood of happening by chance of one in 68,500



Feature	Relationship to Pandemic
Line 2 carried 1 MM passengers a day before COVID	Assuming 2 trips/d for commuters, about 5% of the Wuhan population uses this Line, making it an efficient transmission route for all of Wuhan as well as Hubei Provence. A single patient can leave a droplet/aerosol cloud for hours to infect others.
Line 2 shares stations with every other Metro Line	Permits human-to-human spread to every part of Wuhan at the stations shared with Line 2
Line 2, Hankou Railway Station	Connects Wuhan to all of China by high speed rail
Line 2, Tianhe International Airport	International destinations: New York City, San Francisco, London, Tokyo, Rome, Istanbul, Dubai, Paris, Sydney, Bali, Bangkok, Moscow, Osaka, Seoul, and Singapore.

The Line 2 COVID Conduit

The Line 2 COVID Conduit

The Hunan Seafood Market, Wuhan Institute of Virology, and the Wuhan CDC, all locations suggested to be the possible source of SARS-CoV-2 in Wuhan, are also all serviced by Line 2 of the Metro system, suggesting this public transit line should become the focus for further investigations into the origin of this pandemic.

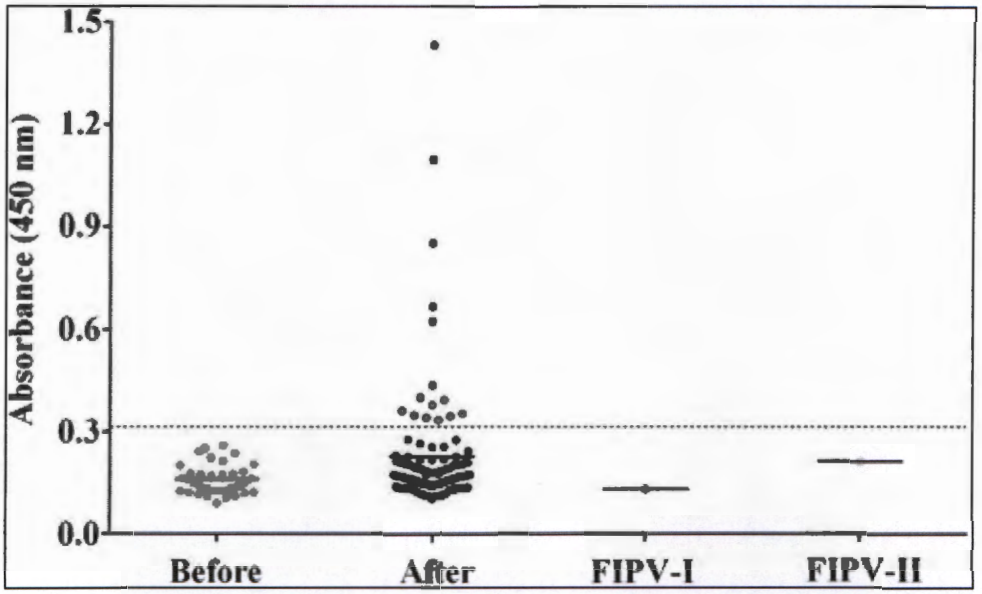
Given that the Hunan Seafood Market has been removed as a source for the origin of CoV-2, this evidence will contribute a 51%/49% contribution in favor of laboratory compared to zoonotic origin. There will be no Subjective Discount Factor adjustment.

The results from the calculations are shown below.

Evidence or process	Zoonotic Origin (ZO)	Laboratory Origin (LO)
Starting likelihood	0.002	0.998
This is the outcome favors LO over ZO at 51% versus 49%		0.51
Impact of this evidence		Increases the likelihood of LO by 51/49 = 1.041
Impact of evidence calculation		1.041 x 0.998 = 1.039
Normalize this step of analysis	0.002/(0.002 + 1.039) = 0.002	1.039/(0.002 + 1.039) = 0.998

Adjusted likelihood: Zoonotic origin (0.2%), laboratory origin (99.8%).

Evidence: **SARS-CoV-2 infection, based on antibody seroconversion, was not found in 39 archived specimens taken from cats (1/3 feral) between March and May 2019[119]**



Based on these results, the prevalence of SARS-CoV-2 in domestic and feral cats prior to January 2020 is less than 8% with a 90% confidence interval.

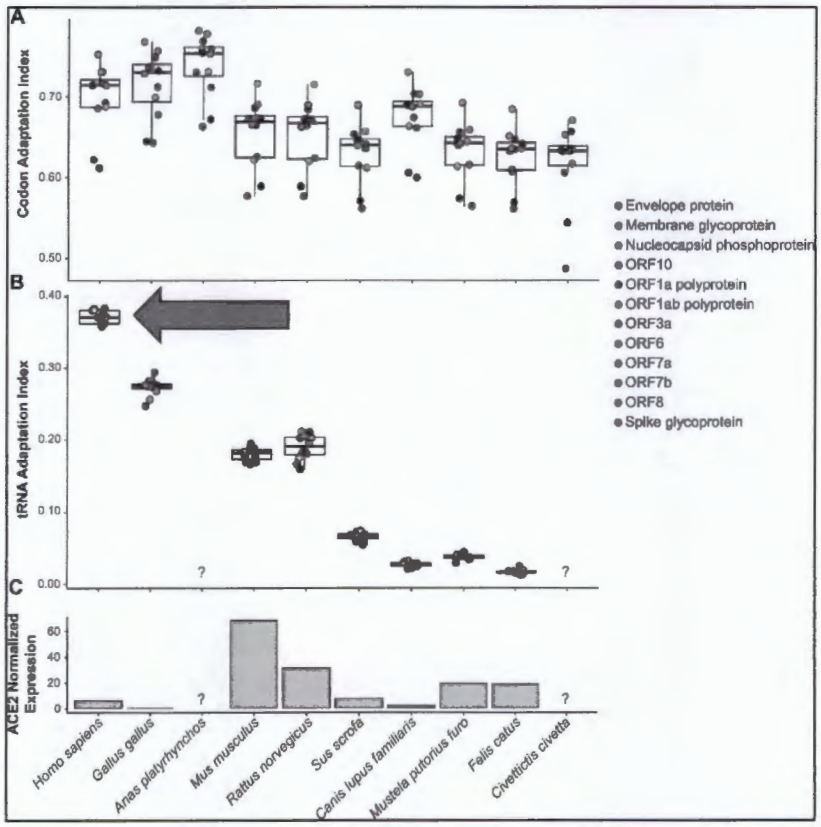
This will not be used to adjust the likelihoods.

Current likelihood: Zoonotic origin (0.2%), laboratory origin (99.8%).

[119] https://www.tandfonline.com/doi/full/10.1080/22221751.2020.1817796

Evidence: The extraordinary pre-adaption of SARS-CoV-2 for human cells is demonstrated by a paper looking at a tRNA adaption index.[120]

"The proteome of SARS-CoV-2 is mainly composed of the replicase polyprotein (ORF1ab) and of structural proteins: the spike glycoprotein, the membrane and envelope proteins, and the nucleoprotein [41]. Based on the genomic codon usage of each of the possible host species, we compute the codon adaptation index (CAI) and the tRNA adaptation index (tAI) to estimate the translational efficiency of SARS-CoV-2 proteins in each host (Fig 3A and 3B and S2 Table). Humans are among the top three species whose CAIs are mostly over 0.70, together with ducks and chickens. In terms of the tAI, humans show the highest translational adaptation among all others, followed by chickens, and, to some extent, mice and rats. On the other hand, cats, ferrets, pigs, and dogs are less translationally adapted than humans both by CAI and tAI."



As shown in panel B above, the tRNA Adaption Index is highest, by far, for humans (blue arrow) followed by the red junglefowl. This is additional evidence of the extraordinary adaption of SARS-CoV-2 to humans from the very beginning. This also is the first evidence of a reasonable intermediate host but based only on these *in silico* data.

This will not be used to adjust the likelihoods.

Current likelihood: Zoonotic origin (0.2%), laboratory origin (99.8%).

[120] https://journals.plos.org/ploscompbiol/article?id=10.1371/journal.pcbi.1008450#pcbi.1008450.s004

Evidence: Evidence of Lax procedures and disregard of laboratory safety protocols and regulations in China, including the Wuhan Institute of Virology

A collection[121] from the Chinese Q&A website, https://www.zhihu.com/, of first-hand documentation of laboratory safety breaches and incidents within a large number of laboratories with diverse research subjects and purposes in the People's Republic of China (PRC) is provided. The laboratories involved include Chemistry labs, Biolabs, Computer labs as well as Physics and Engineering labs.

From this first-hand documentation, we obtained evidence of relaxed safety regulations and frequent breaches of such regulations, with reasons ranging from poor training/education on lab safety and chronic ignorance of safety rules, to intentional breaches of protocols for purposes other than the research projects of the lab(s) of which the breach was documented in.

Such breaches often resulted in safety accidents ranging from physical injury, chemical burns, chemical leaks, and damage to property, to lab-acquired infection and escape of in-lab pathogens. With consequences ranging from personal-level to institution-wide impacts.

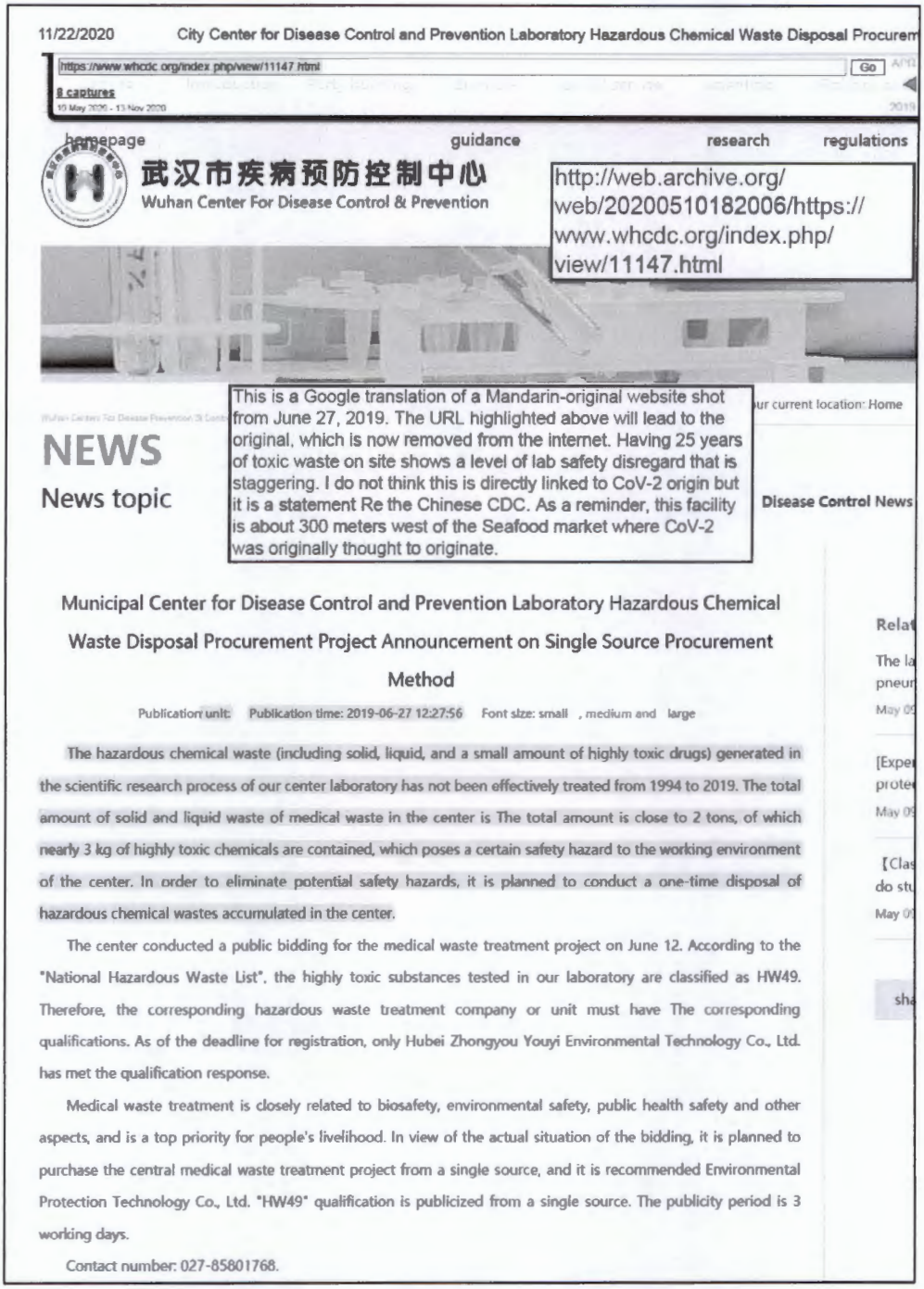
Here is the reference to the State Department cables concerning safety concerns at the WIV.[122]

The following document shows that in June 2019, the Chinese CDC was soliciting for the removal of 25-years-worth of solid and liquid medical waste. The total weight is close to two tons including three kg of highly toxic waste.

This is a Google translation of a Mandarin-original website shot from June 27, 2019. The URL highlighted above will lead to the original, which now has been removed from the internet. Having 25 years of toxic waste on site shows a staggering level of disregard for lab safety.

I do not think this is directly linked to CoV-2 origin, but it is a statement about the Chinese CDC. As a reminder, this facility is about 300 meters west of the Seafood market where CoV-2 was first thought to have originated.

[121] https://zenodo.org/record/4307879#.X-yUo9gzbOh
[122] https://foia.state.gov/Search/Results.aspx?caseNumber=F-2020-05255

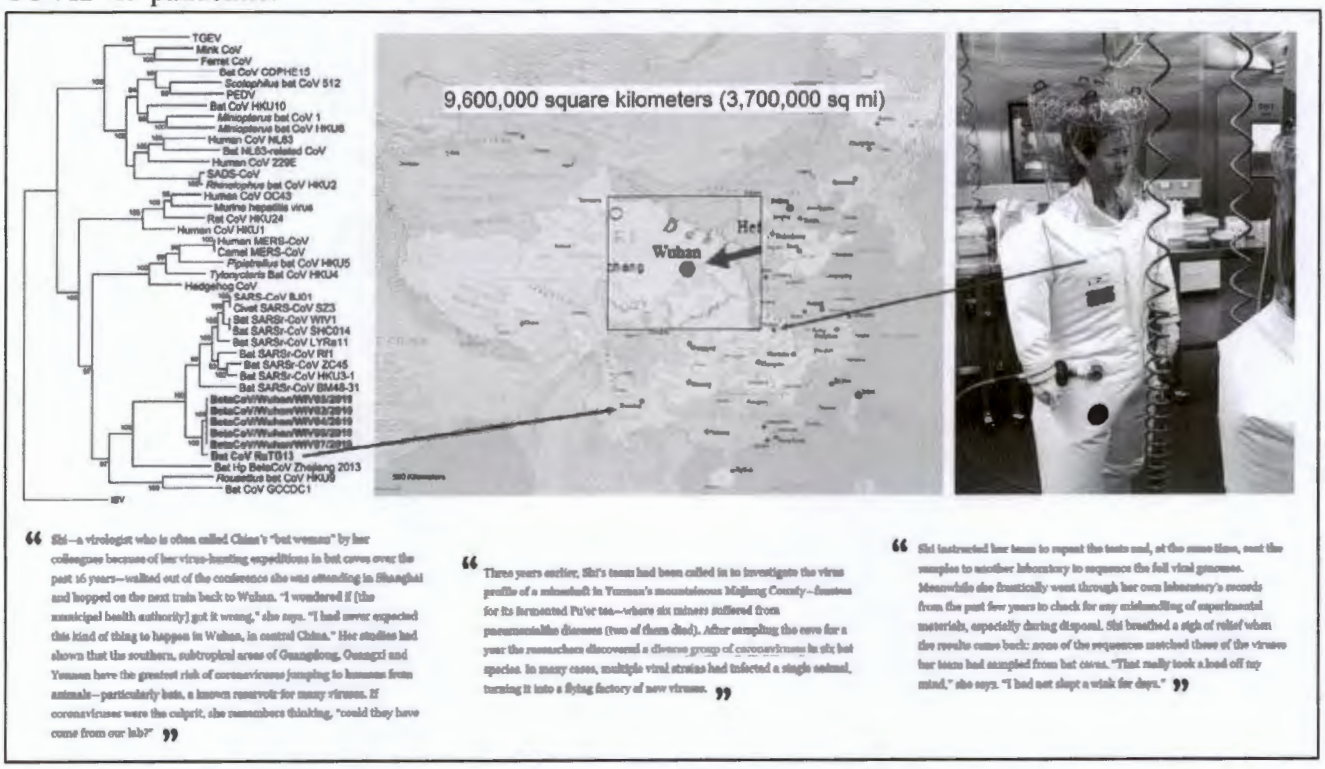


This will not be used to adjust the likelihoods.

Current likelihood: Zoonotic origin (0.2%), laboratory origin (99.8%).

Evidence: The careful words of Dr. Shi do NOT say she did not have SARS-CoV-2 at the WIV.

This Figure contains quotes from an article about Dr. Shi and her reaction to the beginning of the COVID-19 pandemic.



Notice in the last frame Dr. Shi says two strange sentences:

Sentence 1: “…she frantically went through her own laboratory’s records from the past few years to check for any mishandling of experimental materials, **especially during disposal**.”

Why did she mention disposal? If you don’t know what you are looking for this, “especially during disposal,” is a bit of an odd qualifier. Other evidence from Wuhan suggests that, in fact, disposal may have been a likely source of the accidental lab release.

Sentence 2: “She breathed a sigh of relief when the results came back: none of the sequences matched those of the viruses her team had sampled from bat caves.”

If Dr. Shi had created SARS-CoV-2 as a chimera, perhaps starting with one of those cave viruses, of course you would no longer have a sequence match. This is a probably truthful statement that leaves open the question of lab creation.

This will not be used to adjust the likelihoods.

Current likelihood: Zoonotic origin (0.2%), laboratory origin (99.8%).

Evidence: The Good, the Bad, and the Ugly: a review of SARS Lab Escapes[123]

In 2003–04, in the wake of the SARS epidemics, there were multiple cases of laboratory acquired infection (LAI) with SARS within just a few months: first in a P3 in Singapore, then in a military P4 in Taipei and last a protracted case in a P3 in Beijing. The 'WHO SARS Risk Assessment and Preparedness Framework' has a good summary of these lab accidents:

Since July 2003, there have been four occasions when SARS has reappeared. Three of these incidents [note: Singapore, Taipei and Beijing] were attributed to breaches in laboratory biosafety and resulted in one or more cases of SARS. The most recent laboratory incident [note: in Beijing] resulted in 9 cases, 7 of which were associated with one chain of transmission and with hospital spread. Two additional cases at the same laboratory with a history of illness compatible with SARS in February 2004 were detected as part of a survey of contacts at the facility.[i.1]

This article reviews some of these cases and discusses briefly some of the insights that were gained from these at the time.

Another article along the same lines is, "10 incidents discovered at the nation's biolabs"[124] This included Dr. Baric's laboratory in which "(b)etween April 2013 and September 2014, eight individual mouse escapes were reported at the University of North Carolina-Chapel Hill. Several of the mice were infected with either SARS or the H1N1 flu virus."

Dozens of holes in BSL-4 'spacesuits'

As a key protection against the world's most deadly pathogens, including the Ebola virus, scientists in the BSL-4 labs at the U.S. Army Medical Research Institute of Infectious Diseases (USAMRIID) at Fort Detrick in Maryland wear pressurized, full-body spacesuit-like gear and breathe purified air. Yet those suits ruptured or developed holes in at least 37 incidents during a 20-month period in 2013 and 2014, according to lab incident reports obtained by USA TODAY under the federal Freedom of Information Act.

This will contribute a 51%/49% contribution in favor of laboratory compared to zoonotic origin. There will be no confidence adjustment. The results from the calculations are shown below.

Evidence or process	Zoonotic Origin (ZO)	Laboratory Origin (LO)
Starting likelihood	0.011	0.989
The history of SARS laboratory accidents is consistent with the laboratory origin hypothesis		0.51
Impact of this evidence		Increases the likelihood of LO by 51/49 = 1.041
Impact of evidence calculation		1.041 x 0.989 = 1.030
Normalize this step of analysis	0.011/(0.011 + 1.030) = 0.011	1.030/(0.011 + 1.030) = 0.989

Adjusted likelihood: Zoonotic origin (0.2%), laboratory origin (99.8%).

[123] https://gillesdemaneuf.medium.com/the-good-the-bad-and-the-ugly-a-review-of-sars-lab-escapes-898d203d175d

[124] https://www.usatoday.com/story/news/2015/05/29/some-recent-us-lab-incidents/25258237/

Evidence: Drs. Shi and Daszak use Wuhan residents as negative controls for zoonotic coronavirus seroconversion[125]

"As a control, we collected 240 serum samples from random blood donors in **Wuhan >1000 km away from Jinning & where inhabitants have a much lower likelihood of contact with bats due to its urban setting**" [emphasis added]. As expected, 0/240 samples from the patients from Wuhan had a positive serological evidence of prior coronavirus infection.

"The 2.7% seropositivity for the high-risk group of residents living in close proximity to bat colonies suggests that spillover is a **relatively rare event**, however this depends on how long antibodies persist in people, since other individuals may have been exposed and antibodies waned."

In this paper from 2018, Drs. Shi and Daszak conclude that bat-to-human transfer is relatively rare for high-risk people living in close proximity to bat colonies and much less likely in Wuhan, a conclusion that does not support a hypothesis of bat-to-human transmission.

This will not be used to adjust the likelihoods.

Current likelihood: Zoonotic origin (0.2%), laboratory origin (99.8%).

[125] https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6178078/

Evidence. The Bat Coronavirus RaTG13 has the Unique Genome Sequences Necessary to be the Precursor of SARS-CoV-2 Using the 'No See 'Em' Synthetic Biology Technology. ***The probability that RaTG13 acquired these 'No See 'Em' synthetic biology assembly sequences in nature is one in a billion.***

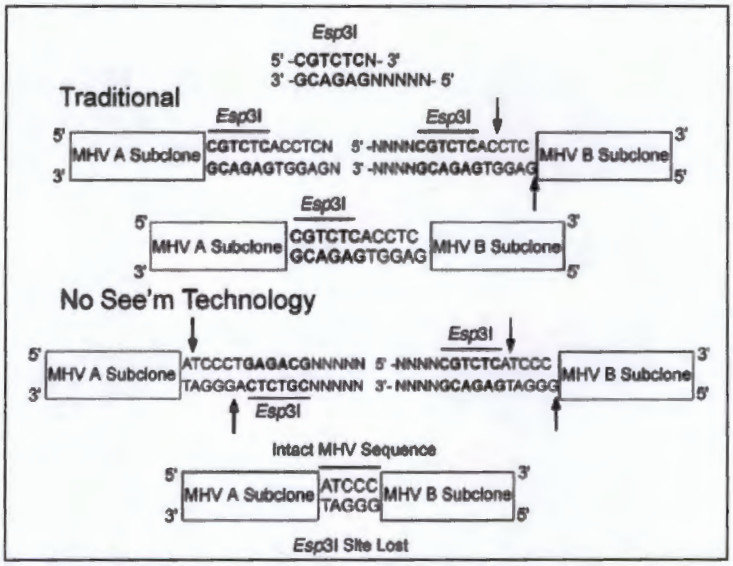
Summary.

- Synthetic biology techniques, like the engineered "No See 'Em'[126] restriction enzyme-enabled insertion method,[127] have been developed that, by design, extinguish the fingerprints of the insertion when only looking at the final genome.

- The use of these techniques is revealed however, if the precursor-product genome pair of such an insertion is available for inspection.

- **Hypothesis: the unique features of the SARS-CoV-2 Spike Protein, the receptor binding domain ACE2 contact amino acid residue region and the polybasic (furin) cleavage site, are the product of a genome insertion sequence into RaTG13 using engineered Esp3I restriction enzyme sites, the so-called, 'No See 'Em,' technology.**

- An example of the 'No See'm' Technology is shown below, taken from Baric and Sim.[1] By placing the restriction sites symmetrically on both strands of the cDNA, the resulting insertion no longer contains the identifying restriction site nts.



- According to Baric and Sims[1] "the type IIS restriction enzyme, Esp3I, recognizes an asymmetric sequence and makes a staggered cut 1 and 5 nucleotides downstream of the recognition sequence, leaving 256, mostly asymmetrical, 4-nucleotide overhangs

[126] Variably spelled 'No See 'Em,' 'No See 'um,' and 'No See'm.'

[127] https://www.researchgate.net/publication/8119695_Development_of_mouse_hepatitis_virus_and_SARS-CoV_infectious_cDNA_constructs

(GCTCTCN#NNNN). As identical Esp3I sites are generated every ~1,000,000 base pairs or so in a random DNA sequence, most restricted fragments usually do not self-assemble."

- Examination of RaTG13 identified two Esp3I cleavage sites in the Spike Protein gene, at nts 1366 and 2941 (positions 22,910 and 24,485 in the entire genome).

- As expected from the above rarity of such sites in an approximately 3800 nt gene, SARS-CoV-2 has no Esp3I sites in its SP gene. Neither do twelve other coronaviruses, including SARS-CoV-1, MERS, and other related human or bat coronaviruses.

- From all of the species other than bat RaTG13 gene source, the frequency of Esp3I sites **at any location** is 2 in 54,131 nucleotides or 0.000036947. If we assume the possibility of the occurrence of such a site at a given nucleotide is independent of any other nucleotide, then it is possible to use a binomial distribution calculation to determine the probability of 2 Esp3I sites in 3809 nucleotides for the bat RaTG13 gene. This calculation yields a probability of at least 2 sites anywhere in the Spike Protein gene of 0.009 or about one in a hundred. The probability of exactly 2 sites is 0.0086.[128]

- The 5' restriction site in RaTG13 begins at aa residue 455L, identified by Andersen et al, Nature, 2020. as the start of the "receptor-binding domain ACE2 contact residues." The downstream amino acids from this site are critical for why RaTG13 has such poor affinity for human ACE2 and the substitutions in CoV-2 are precisely why CoV-2 has such high affinity for human ACE2, why CoV-2 seems so 'preadapted' to human infections, etc. So this is the most important part of CoV-2 in explaining its ACE2 binding and infectivity. Further downstream is arguably the second most important site, the polybasic (furin) cleavage site.[129] Polybasic cleavage sites have not been observed in related 'lineage B' betacoronaviruses,' according to Andersen et al, Nature, 2020. and so there has been much speculation about how this site was acquired.

- The 3' restriction site in RaTG13 is at residue 980L. There is no protein-based rationale for this position.

- Comparing the nt sequences between RaTG13 and CoV-2, at the 5' restriction site, they are two codons in which only 2 of 6 nt bases are shared but, despite this low nt sequence homology, they are in fact synonymous base substitutions.

- Comparing the nt sequence between RaTG13 and CoV-2 at the 3' restriction site, this site has 5 of 6 identical nts with a single synonymous change in CoV-2 which destroys the restriction site. This is the only such five nt site in the RaTG13 spike protein gene and so

[128] Statistical analysis provided by Dr. Martin Lee, PhD, Adjunct Professor of Statistics, UCLA Fielding School of Public Health, UCLA, Los Angeles, CA.
[129] https://www.biorxiv.org/content/10.1101/2020.08.26.268854v1

is the easiest site in which a one nt substitution can create or destroy an Esp3I restriction site.

- The probability of having the restriction sites at **exactly these locations** can also be calculated.[2] Since there are 3809 nucleotides in the RaTG13 genome then, 3807 would not have a restriction site with probability (1-0.000036947), which was determined from the frequency of these restriction sites in other species. The other two sites would have this restriction site with probability 0.000036947. So the overall probability of this configuration has a probability of: $(1\text{-}0.00036947)^{3807} \times (0.000036947)^2 = 3.343 \times 10^{-10}$. This is a frequency of these site at their exact location being here from a natural process of approximately one in a billion.

- Dr. Zhengli-Li Shi, of the Wuhan Institute of Virology, collected the bat virus RaTG13 in 2013 and sequenced it between 2014 and 2018. In 2015, Dr. Shi and colleagues have also used the 'No See 'Em' technology' with a similar restriction enzyme, BgII, in the SARS-CoV reverse genetics system to generate chimeric coronaviruses. In that paper, they inserted a spike protein gene from a bat coronavirus into a mouse-adapted coronavirus, with a 'gain-of-function' phenotypic change.[130]

- **In conclusion:**

- **The bat coronavirus RaTG13 has two rare, Esp3I restriction sites strategically located to permit insertion of a genetic sequence that codes for the unique features of the SARS-CoV-2 Spike Protein, its receptor binding contact amino acids and its polybasic (furin) cleavage site, using the 'No See 'Em' synthetic biology techniques.**

- **This specific synthetic biology laboratory technique has been successfully performed previously by Wuhan Institute of Virology scientists to increase coronavirus infectivity.**

- **The probability these two sites are present and in their exact location in RaTG13 by an act of nature is one in a billion.**

[130] https://www.nature.com/articles/nm.3985

Text-Table. A record of the EspI restriction enzyme sites in the Spike Protein (SP) genes of fifteen coronaviruses, including RaTG13 and SARS-CoV-2. RaTG13 is unique in having two such sites, with SARS-CoV-2 and eleven other coronaviruses having no such site in the SP gene. The restriction sites were identified with the RestrictionMapper site algorithm: http://www.restrictionmapper.org/ .

Species	Spike Protein (SP) Gene Source	Nt Size of SP Gene	Esp3I Site Location in Spike Protein Gene	Reference
Bat	Bat Coronavirus RaTG13 from WIV	3809	1366, 2941 (22910, 24485 in genome)	
Human	SARS-CoV-2 Reference Sequence	3821	None	
Bat	Rhinolophus affinis coronavirus isolate LYRa11	3779	None	Daszak and Shi paper
Bat	Bat SARS coronavirus HKU3-1	3728	None	Daszak and Shi paper
Bat	SARS-like coronavirus isolate bat-SL-CoVZC45	3740	None	Third Military University publication
Bat	SARS-like coronavirus bat-SL-CoVZXC21	3737	None	Third Military University publication
Bat	hCoV-19/bat/Yunnan/RmYN02/2019	3873	None	Wild bat coronavirus with apparent furin-like insert
Bovine	Bovine coronavirus strain Quebec	4091	None	
Human	Human coronavirus HKU1 strain	4070	3208	
Human	MERS Reference Sequence	4061	None	
Human	Human coronavirus OC43 strain	4079	None	
Human	Human coronavirus 229E strain	3512	None	
Human	Human Coronavirus NL63 Reference Sequence	4070	None	
Human	SARS 2003 coronavirus ZJ0301	3767	None	
Pangolin	Pangolin coronavirus isolate PCoV GX-P4L	3803	3351	
Human	SARS-CoV-1 Urbani	3767	None	

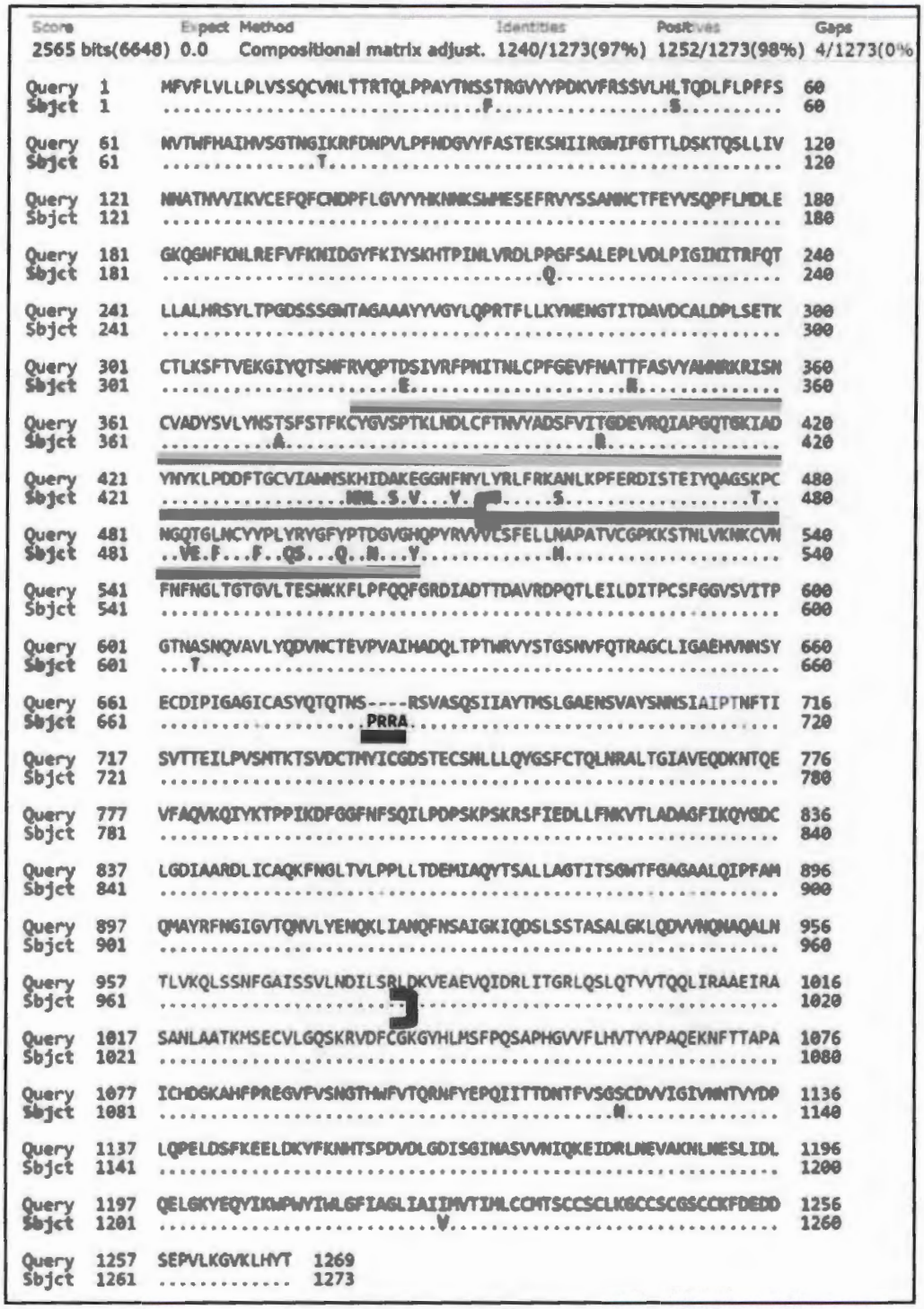
Figure. A comparison of the RaTG13 Spike Protein gene (Query) and the SARS-CoV-2 Reference Sequence (Sbjct) showing the only two Esp3I restriction enzyme cleavage site, both present in RaTG13 but absent in SARS-CoV-2. The restriction sites were identified with the RestrictionMapper site: http://www.restrictionmapper.org/ .The 5' cleavage site is strategically located at the beginning of the receptor binding domain ACE2 contact residues. Despite four of six nt are different these are synonymous changes.

```
Query  1321  ATTGATGCAAAAGAGGGCGGTAATTTTAACTATCTTTACCGTCTCTTTAGAAAAGCTAAT  1380
Sbjct  1321  CTTGATTCTAAGGTTGGTGGTAATTATAATTACCTGTATAGATTGTTTAGGAAGTCTAAT  1380
```

The 3' cleavage site is the only downstream -CGTCTN- sequence found in the CoV-2 Spike Protein, making it unique.

```
Query  2927  TCCTTTCACGTCTCGACAAAGTTGAGGCTGAAGTGCAGATTGACAGGTTGATCACAGGCA  2986
Sbjct  2939  TCCTTTCACGTCTTGACAAAGTTGAGGCTGAAGTGCAAATTGATAGGTTGATCACAGGCA  2998
```

Figure. Comparison of Spike Protein amino acid sequence between RaTG13 (Query) and SARS-CoV-2 (Sbjct). Amino acid substitutions in CoV-2 are shown in red, single letter abbreviation. Green band; receptor binding domain. Blue band; receptor binding domain ACE2 contact residues (Andersen et al, Nature, 2020.). Purple band; polybasic (furin) cleavage site. Red brackets; Esp3I cleavage sites in RaTG13.



Because it has not been established that RaTG13 was the precursor of CoV-2 this evidence statement will not be used at this time to adjust the likelihoods of the origin. If additional information is obtained at a later date this may be revisited.

Likelihood from prior state is unchanged following this evidence analysis:

Zoonotic origin (0.2%) and laboratory origin (99.8%)

Evidence. Location, location, location: Based on the distance between known SARS-CoV-1 laboratory-acquired infections and the hospital of admission of the infected personnel, the WIV is within the expected hospital catchment for a CoV-2 LAI

Hypothesis. Laboratory-acquired infections (LAI) have the property that the hospital of admission of the personnel from the laboratory with the acquired infection are close together, specifically they are within 24.64 km from the laboratory.

Prior data from SARS-CoV-1. There were four LAIs of SARS-CoV-1 that can be used to determine the distance between the laboratory where the infection occurred and the hospital of first admission. The data are here:

SARS-CoV-1 Laboratory Acquired Infection (LAI)	Hospital of admission	Distance (Google Maps)	
In September 2003, a 27-year-old student from the National University of Singapore (NUS) was infected with the SARS virus due to improper experimental procedures	Singapore General Hospital (SGH)	6.3 km	
Baiji Mountain, Sanxia, Taiwan	Taiwan Hoping Hospital, Taipei, Taiwan	27.8 km	
№100 Yingxin Street, Xicheng District, Bejing	Union Hospital, Beiijing, China	7.3 km	
№100 Yingxin Street, Xicheng District, Bejing	Friendship Hospital, Beijing, China	17.6 km	
		mean = 14.75	
		SD = 10.1	
		95% Confidence Interval	14.75 ±9.887

Based on these four cases, the 95% upper confidence limit for the distance from LAI patients to the hospitals of admission is 24.6 km of the laboratory where the infection was acquired.

SARS-CoV-2. Although it is not clear which hospital the first patient was admitted to the following Text-Table contains all likely candidates.

SARS-CoV-2 Potential LAI Source	Hospital of admission	Distance (Google Maps)	Probability of being closer than the average results for SARS-CoV-1	Probability of being farther than the average results for SARS-CoV-1
Wuhan Institute of Virology, Wuhan, China	PLA Hospital, NO. 627 Wuluo Road, Wuchang District, Wuhan, China	4.8 km	0.094	0.906
Wuhan Institute of Virology, Wuhan, China	Wuhan Central Hospital, Wuhan, China	9.1 km	0.338	0.662
Wuhan Institute of Virology, Wuhan, China	Zhongnan Hospital, Wuhan, China	2.8 km	0.019	0.981
Wuhan Institute of Virology, Wuhan, China	Tongji Hospital, Wuhan, China	5.1 km	0.109	0.891
Wuhan Institute of Virology, Wuhan, China	Hubei Maternity and Child Health Care Hospital, Wuhan, China	4.4 km	0.075	0.925
Hypothesis: Given the distance from the SARS-CoV-1 laboratory where an LAI occurred to the hospital of admission for the lab workers who became infected, what is the probability that CoV-2 is also an LAI, given the distance from the hospitals where the first patients were seen to the WIV, the hypothesized source.			Probability calculations based on the use of a log-normal distribution for distances	Probability calculations based on the use of a log-normal distribution for distances

Based on the data for actual LAI for SARS-CoV-1 the distance between the WIV and the hospitals of admission for CoV-2 is consistent with the WIV being the origin for the LAI. There is no evidence the putative LAI for CoV-2 is any different than the known LAIs for CoV-1.

This evidence is not independent of other evidence that is based on location and so it cannot be used independently in the Bayesian analysis. It is included here for completeness.

Likelihood from prior state is unchanged following this evidence analysis:

Zoonotic origin (0.2%) and laboratory origin (99.8%)

Evidence. Dr. Shi successfully identifies a laboratory-acquired infection outbreak from Hanta virus in laboratory rodents.

Infection, Genetics and Evolution 10 (2010) 638–644

Contents lists available at ScienceDirect

Infection, Genetics and Evolution

journal homepage: www.elsevier.com/locate/meegid

Hantavirus outbreak associated with laboratory rats in Yunnan, China

Yunzhi Zhang [a,b], Hailin Zhang [b,*], Xingqi Dong [b], Junfa Yuan [a], Huajun Zhang [a], Xinglou Yang [a], Peng Zhou [a], Xingyi Ge [a], Yan Li [a], Lin-Fa Wang [c], Zhengli Shi [a,**]

[a] State Key Laboratory of Virology, Wuhan Institute of Virology, Chinese Academy of Sciences, Wuhan, People's Republic of China
[b] Yunnan Institute of Endemic Diseases Control and Prevention, Dali, People's Republic of China
[c] Commonwealth Scientific and Industrial Research Organisation Livestock Industries, Geelong, Victoria, Australia

ARTICLE INFO

Article history:
Received 16 November 2009
Received in revised form 20 February 2010
Accepted 30 March 2010
Available online 7 April 2010

Keywords:
Hemorrhagic fever with renal syndrome
Hantavirus
Laboratory rats
Recombination

ABSTRACT

An outbreak of hemorrhagic fever with renal syndrome occurred among students in a college (College A) in Kunming, Yunnan province, China in 2003. Subsequent investigations revealed the presence of hantavirus antibodies and antigens in laboratory rats at College A and two other institutions. Hantavirus antibodies were detected in 15 additional individuals other than the index case in these three locations. Epidemiologic data indicated that the human infections were a result of zoonotic transmission of the virus from laboratory rats. A virus was isolated from rats in College A and the full-length genome sequence revealed that this was a new Hantaan virus isolate, designated strain KY. Sequence analysis of the three genome segments indicated that this new isolate is a reassortant derived from human and rat Hantaan viruses. Further sequence analysis of the medium (M) genome segment revealed that it originated from a recombination event between two rat Hantaan virus lineages.

The significance of this evidence is that it demonstrates the methods used by Dr. Shi and the WIV to solve a laboratory-acquired infection outbreak. The methods described herein should be applied to the WIV in order to determine if CoV-2 was also a laboratory-acquired infection.

This will not be used to directly advance the Bayesian analysis.

Likelihood from prior state is unchanged following this evidence analysis:

Zoonotic origin (0.2%) and laboratory origin (99.8%)

Evidence. Bats hibernate when the temperature is below 10.5 C;[131] in Hubei province that begins in September and ends in May.

Average Hubei Temperature by Month

Month	Recommended Rate	Max Temp.		Min Temp.	
Jan.	~		-17°C		-26°C
Feb.	√		-13°C		-23°C
Mar.	√√		-3°C		-14°C
Apr.	√√		7°C		-4°C
May.	√√		16°C		4°C
Jun.	√		23°C		11°C
Jul.	√		23°C		13°C
Aug.	√		21°C		11°C
Sep.	√√		15°C		4°C
Oct.	√√		6°C		-5°C
Nov.	√√		-5°C		-16°C
Dec.	√		-15°C		-23°C

Based on this evidence, they would have been hibernating at the time of the first human outbreak in the fall of 2019. Since this evidence is cumulative to the prior evidence from Dr. Shi that the bat host species for CoV-2 does not live in Hubei Province it will not be used to change the Bayesian analysis.

Likelihood from prior state is unchanged following this evidence analysis:

Zoonotic origin (0.2%) and laboratory origin (99.8%)

[131] https://zslpublications.onlinelibrary.wiley.com/doi/abs/10.1111/j.1469-7998.1971.tb01323.x

Wuhan Institute of Virology analysis of lavage specimens from ICU patients at Wuhan Jinyintan Hospital in December 2019 contain both SARS-CoV-2 and adenovirus vaccine sequences consistent with a vaccine challenge trial

Summary. The most significant evidence provided herein is the finding from RNA-Seq performed by the Wuhan Institute of Virology (WIV) of lavage patient samples collected on December 30, 2019.[132] These ICU patients were the subject of the seminal paper, entitled, "A pneumonia outbreak associated with a new coronavirus of probable bat origin," from Dr. Zhengli Shi and colleagues that first characterized SARS-CoV-2.[133] This author has confirmed that the RNA-Seq of all five patients contained SARS-CoV-2 sequences.

Surprisingly the specimens also contained the adenovirus "pShuttle" vector, developed by Chinese scientists in 2005 for SARS-CoV-1.[134] Two immunogens were identified, the Spike Protein gene of SARS-CoV-2 and the synthetic construct H7N9 HA gene.[135] Hundreds of perfectly homologous (150/150) raw reads suggest this is not an artifact. Reads that cross the vector-immunogen junction are identified. While adenovirus is a common infection the wildtype viruses have low homology to the vaccine vector sequence, by design, to avoid rejection of the vaccine due to prior exposure to wildtype adenoviruses.

Two patients from the same hospital who had bronchial lavage on the same day but had their specimens sent to the Hubei CDC did not have adenovirus vaccine sequences.

Three explanations come to mind from this evidence:

1. These represent sample preparation artifacts at the WIV, such as sample spillover on the sequencer.
2. These patients were admitted with an unknown infection, were not responding to the treatment protocols for a infection of unknown origin, and they were vaccinated with an experimental vaccine in a desperate but compassionate therapeutic "Hail Mary."
3. A clinical trial of a combination[136] influenza/SARS-CoV-2 vaccine was being conducted and an accidental release into Wuhan occurred.

Only WIV scientists and Chinese authorities can answer these questions. Until the evidence of the adenovirus sequences has been confirmed by other scientists, this author will not include this evidence in the Bayesian analysis.

Obviously if a vaccine containing the Spike Protein of SARS-CoV-2 was being administered to patients in Wuhan in December 2019 the question of laboratory origin is a settled matter.

---

[132] The detailed evidence for the adenovirus vaccine sequences is given at the end of this document.
[133] https://www.nature.com/articles/s41586-020-2012-7
[134] https://www.ncbi.nlm.nih.gov/nuccore/AY862402.1
[135] https://www.ncbi.nlm.nih.gov/nuccore/K[illegible]425.1/
[136] The proposal that this was, in fact, a combination vaccine was made by H. Lawrence Remmel, Department of Pathology, University Medical Center Utrecht, Utrecht University, Utrecht, The Netherlands.

Introduction. Following the 2003 SARS epidemic, Liu et al. developed an adenoviral expression vector of a truncated S1 subunit of SARS-CoV spike protein that resulted in specific humoral immune responses against SARS-CoV in rats.[137] This same vector was used to create the CoV-2 adenovirus vector vaccine.

In order to test the hypothesis that CoV-2 began in the PLA Hospital as a vaccine challenge clinical trial that went awry, RNA-Seq raw reads from nasopharyngeal specimens of Wuhan COVID patients (Table below) were blasted against the published genome sequence of the SARS-CoV-1 vaccine (GenBank AY862402.1). I used the SARS-CoV-1 vaccine because the PLA CoV-2 vaccine has not been published at this time.

Adenovirus sequences detected	GenBank URL	GenBank Biosample URL	GISAID ID	CoV-2 Isolate	Sequencing Institution	Clinical Information from GISAID
>100	SRX7730879	SAMN14082200	EPI_ISL_402130	WIV07; Lineage B; mutations NSP3 D1761A, NSP4 T327I; passage original	Wuhan Institute of Virology, Chinese Academy of Sciences	56 y, male, hospitalized, ICU10G, 20 Dec 2019
>100	SRX7730880	SAMN14082196	EPI_ISL_402127	WIV02; Lineage B; mutations NSP16 D220N; passage original	Wuhan Institute of Virology, Chinese Academy of Sciences	32 y, male, hospitalized, ICU4G, outbreak 19 Dec 2019
>100	SRX7730881	SAMN14082197	EPI_ISL_402124	WIV04; Lineage B; no mutations; passage original	Wuhan Institute of Virology, Chinese Academy of Sciences	49 y, female, hopitalized, ICU-6, outbreak 27 Dec 2019, Retailer at Huanan Seafood Wholesale Market, patient alive
>100	SRX7730882	SAMN14082198	EPI_ISL_402128	WIV05; Lineage B; NSP3 G1433S, NSP16 K160R; passage original	Wuhan Institute of Virology, Chinese Academy of Sciences	52 y, female, hospitalized, ICU8G, outbreak 22 Dec 2019; recovered
>100	SRX7730883	SAMN14082199	EPI_ISL_402129	WIV06; Lineage B; no mutations; original passage	Wuhan Institute of Virology, Chinese Academy of Sciences	40 y, male, hospitalized, ICU9G, 25 Dec 2019
>100	SRX7730884	SAMN14082200	EPI_ISL_402130	WIV07; Lineage B; mutations NSP3 D1761A, NSP4 T327I; passage original	Wuhan Institute of Virology, Chinese Academy of Sciences	56 y, male, hospitalized, ICU10G, 20 Dec 2019
7 small	SRX7730885	SAMN14082196	EPI_ISL_402127	WIV02; Lineage B; mutations NSP16 D220N	Wuhan Institute of Virology, Chinese Academy of Sciences	32 y, male, hospitalized, ICU, outbreak 19 Dec 2019
1 small one	SRX7730886	SAMN14082197	EPI_ISL_402124	WIV04; Lineage B; no mutations; passage original	Wuhan Institute of Virology, Chinese Academy of Sciences	49 y, female, hopitalized, ICU-6, outbreak 27 Dec 2019, Retailer at Huanan Seafood Wholesale Market, patient alive
Very few	SRX7730887	SAMN14082199	EPI_ISL_402129	WIV06; Lineage B; no mutations; original passage	Wuhan Institute of Virology, Chinese Academy of Sciences	40 y, male, hospitalized, ICU9G, 25 Dec 2019
None	SRX8032202	SAMN14479127	EPI_ISL_412898	hCoV-19/Wuhan/HBCDC-HB-02/2019	Hubei Provincial Center for Disease Control and Prevention	male, "traveled from Wuhan"
None	SRX8032203	SAMN14479128	EPI_ISL_402132	Wuhan HBCDC-HB-01/2019; Lineage B; mutation Spike F32I; original passage	Hubei Provincial Center for Disease Control and Prevention	49 y, female, hospitalized

This is not related to the previous claim, now shown to be wrong, that SARS-CoV-2 itself contained adenovirus pShuttle sequences.[138]

[137] https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7114075/
[138] https://sciencefeedback.co/claimreview/2019-novel-coronavirus-2019-ncov-does-not-contain-pshuttle-sn-sequence-no-evidence-that-virus-is-man-made/